M208

Pure mathematics

Book F

Analysis 2

This publication forms part of an Open University module. Details of this and other Open University modules can be obtained from Student Recruitment, The Open University, PO Box 197, Milton Keynes MK7 6BJ, United Kingdom (tel. +44 (0)300 303 5303; email general-enquiries@open.ac.uk).

Alternatively, you may visit the Open University website at www.open.ac.uk where you can learn more about the wide range of modules and packs offered at all levels by The Open University.

The Open University, Walton Hall, Milton Keynes, MK7 6AA.

First published 2018.

Edited, designed and typeset by The Open University, using LaTeX.

Printed in the United Kingdom by Hobbs the Printers Limited, Brunel Road, Totton, Hampshire, SO40 3WX.

ISBN 978 1 4730 2349 9

2.1

Contents

Unit F1

Limits

Introduction to Book F

In this book we put the foundations of calculus on a firm logical basis. At the end of Book D *Analysis 1* you met the idea of a continuous function, and at the start of this book you will meet the related ideas of a *limit* of a function and of *uniform continuity*. Then we use these ideas to study *differentiation* and *integration* in detail. Finally, we discuss the representation of functions by *power series*.

You will meet many applications of these ideas, including:

- a technique for proving inequalities such as

  $$\log(1+x) > x - \tfrac{1}{2}x^2, \quad \text{for } x \in (0, \infty)$$

- a function which is continuous but nowhere differentiable
- Stirling's approximate formula for $n!$
- several remarkable exact formulas for π, including Wallis' Product,

  $$\frac{\pi}{2} = \lim_{n\to\infty} \left(\frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5} \cdot \frac{6}{5} \cdot \frac{6}{7} \cdot \cdots \cdot \frac{2n}{2n-1} \cdot \frac{2n}{2n+1} \right)$$

- an elegant proof that π is irrational.

Introduction

In this unit you will meet the concept of a *limit of a real function*, which is closely related to the idea of a continuous function. Roughly speaking, a real function f has a limit at a point c if it is either continuous at c, or if it is defined near c and we can assign a value to $f(c)$ that makes the function continuous at c. You will also study various types of asymptotic behaviour of functions – that is, their behaviour when the domain variable or the codomain variable becomes arbitrarily large. For example, you will see that if $n \in \mathbb{N}$, then $x^n/e^x \to 0$ as $x \to \infty$.

In the second half of the unit you will return to the topic of continuity and study an alternative definition, the so-called ε-δ definition of continuity. This is equivalent to the definition based on sequences that you studied in Unit D4 *Continuity* and, although it may appear to be more abstract, it is easier to use in certain situations. You will meet several unusual functions and see how the two definitions of continuity can be used to investigate at which points they are continuous or discontinuous. You may be surprised at the results!

Finally, you will meet the concept of *uniform continuity*. This is a stronger form of continuity, defined using the ε-δ approach, and it will play an important role when you study the integration of continuous functions.

Many of the results in Sections 1 and 2 are analogues of results on sequences and continuity covered in previous units, so we omit their proofs.

1 Limits of functions

In this section you will study the concept of a limit of a real function (that is, a function whose domain and codomain are subsets of $\mathbb{R}$). The notion of a limit is of fundamental importance to differentiation, which you will study in the next unit, so you should make sure you have a good understanding of this material.

1.1 What is a limit of a function?

Sometimes we need to understand the behaviour of a function that is defined near a particular point, but not at the point itself. For example, consider the function

$$f(x) = \frac{\sin x}{x} \quad (x \in \mathbb{R} - \{0\}).$$

(This function arises when we prove that the sine function is differentiable, as you will see in Unit F2 *Differentiation*.) The graph of f in Figure 1 suggests that as x gets closer and closer to 0, $f(x)$ takes values which are closer and closer to 1.

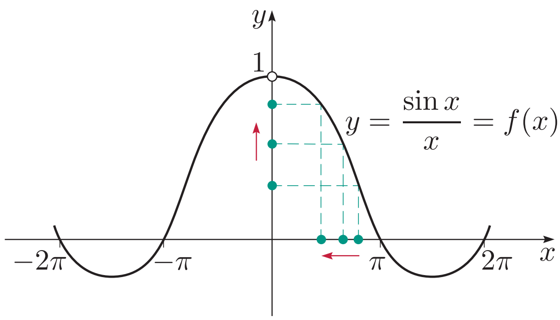


Figure 1 The graph of $y = \dfrac{\sin x}{x}$

On the other hand, consider the function

$$g(x) = \sin\frac{1}{x} \quad (x \in \mathbb{R} - \{0\}).$$

In this case, when x takes values close to 0, the values taken by $g(x)$ do not lie close to any *single* real number: as you can see in Figure 2, g is highly oscillatory near 0.

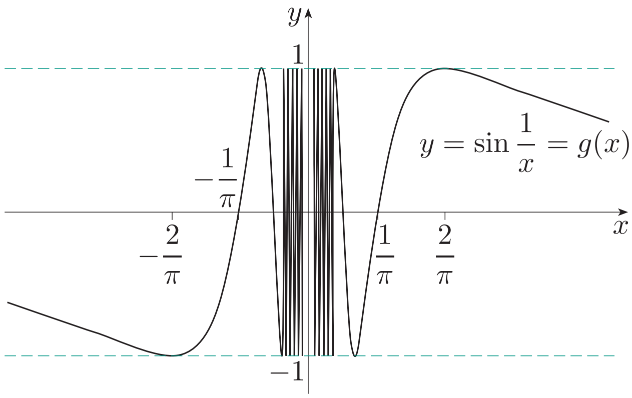


Figure 2 The graph of $y = \sin\frac{1}{x}$

We say that the function f has a *limit* as x tends to 0, but the function g does not. We now make this concept precise.

First we introduce the idea of a **punctured neighbourhood** of a point c. This is simply a bounded open interval with midpoint c, from which the point c itself has been removed. We use the notation $N_r(c)$ for a punctured neighbourhood of length $2r$ with centre c, so

$$N_r(c) = (c - r, c) \cup (c, c + r), \quad \text{where } r > 0,$$

as illustrated in Figure 3. For example $N_1(3) = (2, 3) \cup (3, 4)$.

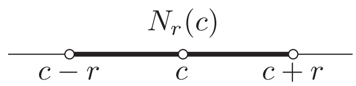


Figure 3 The punctured neighbourhood $N_r(c)$

Definition

Let f be a function defined on a punctured neighbourhood $N_r(c)$ of c. Then $f(x)$ **tends to the limit l as x tends to c** if $l \in \mathbb{R}$ and

for each sequence (x_n) in $N_r(c)$ such that $x_n \to c$,
$f(x_n) \to l$.

In this case, we write

$$\lim_{x \to c} f(x) = l \quad \text{or} \quad f(x) \to l \text{ as } x \to c.$$

Remarks

1. The definition applies whether or not the function f is defined at the point c and, if it is defined there, irrespective of the value of $f(c)$.
2. Note that the limit $\lim_{x \to c} f(x)$, if it exists, does not depend on which punctured neighbourhood of c is considered (that is, it does not depend on r). It is also important to note that the limit must be the same for *every* possible sequence (x_n).

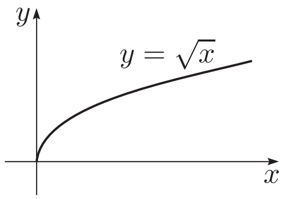


Figure 4 The graph of $y = \sqrt{x}$

3. The above definition does not allow us to state that $\lim_{x\to 0} \sqrt{x}$ exists, because the domain $[0,\infty)$ of $f(x) = \sqrt{x}$ does not contain any punctured neighbourhood of 0. Later in this section we introduce the idea of a *one-sided limit*, and see that $f(x) = \sqrt{x}$ has limit 0 as x tends to 0 *from the right*. The graph of this function is shown in Figure 4.
4. Because this definition of a limit involves sequences, we often use results about sequences to determine whether a function has a limit at a particular point.

We now use this definition to prove that $(\sin x)/x \to 1$ as $x \to 0$, as we guessed earlier.

Theorem F1

$$\lim_{x\to 0} \frac{\sin x}{x} = 1.$$

Proof To prove this result we first use two trigonometric inequalities to establish upper and lower bounds for $(\sin x)/x$ on a punctured neighbourhood of 0. This then enables us to apply the Squeeze Rule for convergent sequences from Subsection 3.3 of Unit D2 *Sequences*.

First note that the function $x \longmapsto (\sin x)/x$ is defined on every punctured neighbourhood of 0.

We now use the inequality

$$\sin x \leq x, \quad \text{for } 0 < x \leq \pi/2,$$

(proved in Subsection 2.3 of Unit D4) to deduce that

$$\frac{\sin x}{x} \leq 1, \quad \text{for } 0 < x \leq \pi/2. \qquad (1)$$

Next we require the inequality

$$x \leq \tan x, \quad \text{for } 0 < x < \pi/2, \qquad (2)$$

which follows by comparing the area of a sector of a disc of radius 1 (shown in Figure 5(a)) with that of a certain right-angled triangle which contains the sector (shown in Figure 5(b)).

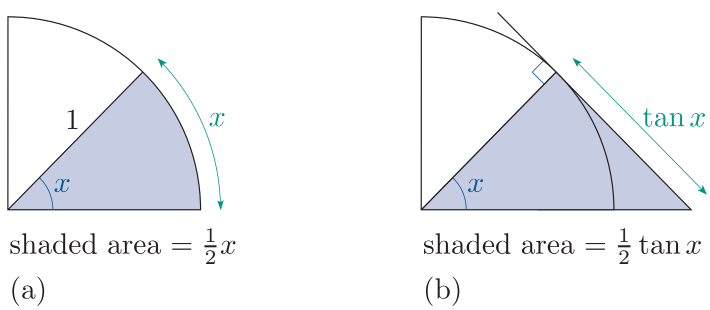


Figure 5 (a) A sector of a disc (b) A triangle containing the sector

Recall that a sector of angle θ in a disc of radius r has area $\frac{1}{2}\theta r^2$.

Since $\cos x > 0$ for $0 < x < \pi/2$, we deduce from inequality (2) that

$$\cos x \leq \frac{\sin x}{x}, \quad \text{for } 0 < x < \pi/2.$$

Remember that a real function f is *even* if $f(x) = f(-x)$ for $x \in \mathbb{R}$.

Thus, by inequality (1) and the fact that the functions $x \longmapsto \cos x$ and $x \longmapsto (\sin x)/x$ are both even,

$$\cos x \leq \frac{\sin x}{x} \leq 1, \quad \text{for } 0 < |x| < \pi/2. \tag{3}$$

We have now established lower and upper bounds for $(\sin x)/x$ in the punctured neighbourhood $N_{\pi/2}(0)$.

Now suppose that (x_n) is any null sequence in the punctured neighbourhood $N_{\pi/2}(0)$. Then

$$\cos x_n \leq \frac{\sin x_n}{x_n} \leq 1, \quad \text{for } n = 1, 2, \ldots, \tag{4}$$

by inequalities (3). Since $x_n \to 0$, we have $\cos x_n \to 1$, because the cosine function is continuous at 0 and $\cos 0 = 1$. Hence, by inequalities (4) and the Squeeze Rule for sequences,

$$\frac{\sin x_n}{x_n} \to 1 \text{ as } n \to \infty,$$

as required. ■

The limit in Theorem F1 was quite tricky to establish, but usually there are simpler ways to find limits. For example, we can determine many limits of functions by using the Combination Rules for sequences which you met in Unit D2.

Worked Exercise F1

Prove that each of the following functions tends to a limit as x tends to 2, and determine these limits.

(a) $f(x) = \dfrac{x^2 - 4}{x - 2}$ (b) $f(x) = \dfrac{x^3 - 3x - 2}{x^2 - 3x + 2}$

Solution

(a) The domain of f is $\mathbb{R} - \{2\}$, so f is defined on each punctured neighbourhood of 2. Also,

$$f(x) = \frac{x^2 - 4}{x - 2} = \frac{(x-2)(x+2)}{x - 2} = x + 2, \quad \text{for } x \neq 2.$$

We can cancel $x - 2$, since $x \neq 2$.

Thus if (x_n) is any sequence in $\mathbb{R} - \{2\}$ such that $x_n \to 2$, then

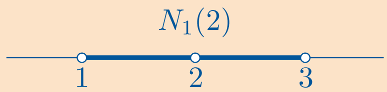
$$f(x_n) = x_n + 2 \to 2 + 2 = 4 \quad \text{as } n \to \infty,$$

by the Sum Rule for sequences. Hence

$$\lim_{x \to 2} \frac{x^2 - 4}{x - 2} = 4.$$

(b) Since $x^2 - 3x + 2 = (x-2)(x-1)$, the domain of f is $\mathbb{R} - \{1, 2\}$.

Because f is not defined at 1, the largest punctured neighbourhood of 2 in which f is defined is $N_1(2)$, illustrated below.



Thus f is defined on $N_1(2)$ and

$$f(x) = \frac{x^3 - 3x - 2}{x^2 - 3x + 2} = \frac{(x-2)(x^2 + 2x + 1)}{(x-2)(x-1)} = \frac{x^2 + 2x + 1}{x - 1},$$

for $x \in N_1(2)$. Thus if (x_n) lies in $N_1(2)$ and $x_n \to 2$, then

$$f(x_n) = \frac{x_n^2 + 2x_n + 1}{x_n - 1} \to \frac{4 + 4 + 1}{2 - 1} = 9,$$

by the Combination Rules for sequences. Hence

$$\lim_{x \to 2} \frac{x^3 - 3x - 2}{x^2 - 3x + 2} = 9.$$

Later in this section you will meet further techniques for finding limits. First, however, we give a strategy for proving that a limit does *not* exist.

Strategy F1

Let f be a real function defined on a punctured neighbourhood $N_r(c)$ of c.

To show that $\lim_{x \to c} f(x)$ does not exist, either

- find two sequences (x_n) and (y_n) in $N_r(c)$ which tend to c, such that $(f(x_n))$ and $(f(y_n))$ have different limits, or
- find a sequence (x_n) in $N_r(c)$ which tends to c such that $f(x_n) \to \infty$ or $f(x_n) \to -\infty$.

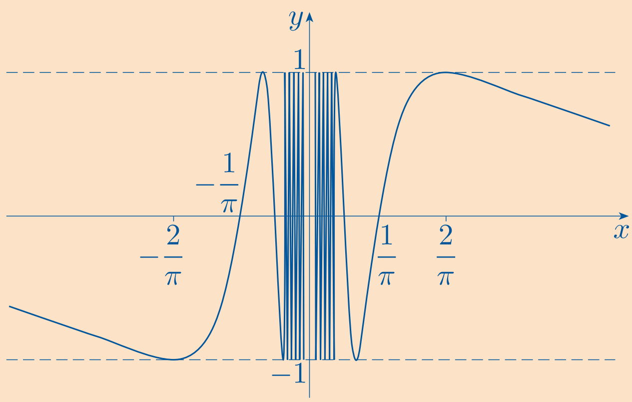
Worked Exercise F2

Prove that each of the following functions does not tend to a limit as x tends to 0.

(a) $f(x) = \sin(1/x)$ (b) $f(x) = 1/x$

Solution

(a) The graph of f is shown below.



The graph suggests that we should be able to use the first part of Strategy F1, with one sequence of points whose images under f are equal to 1 and one sequence of points whose images under f are equal to -1.

The function $f(x) = \sin(1/x)$ has domain $\mathbb{R} - \{0\}$. To prove that $f(x)$ does not tend to a limit as x tends to 0, we choose two null sequences (x_n) and (y_n) in $\mathbb{R} - \{0\}$ such that

$$f(x_n) \to 1 \quad \text{whereas} \quad f(y_n) \to -1.$$

To do this we use the facts that

$$\sin(2n\pi + \pi/2) = 1 \quad \text{and} \quad \sin(2n\pi + 3\pi/2) = -1, \quad \text{for } n \in \mathbb{Z}.$$

It follows that if we choose

$$x_n = \frac{1}{2n\pi + \pi/2} \quad \text{and} \quad y_n = \frac{1}{2n\pi + 3\pi/2}, \quad n = 1, 2, \ldots,$$

then $x_n \to 0$, $y_n \to 0$ and, for $n = 1, 2, \ldots$,

$$f(x_n) = \sin(1/x_n) = 1 \quad \text{and} \quad f(y_n) = \sin(1/y_n) = -1.$$

So

$$f(x_n) \to 1 \text{ as } n \to \infty \quad \text{and} \quad f(y_n) \to -1 \text{ as } n \to \infty.$$

Hence $f(x) = \sin(1/x)$ does not tend to a limit as x tends to 0.

(b) The behaviour of $f(x)$ as $x \to 0$ suggests that we can use the second part of Strategy F1.

The function $f(x) = 1/x$ has domain $\mathbb{R} - \{0\}$. The sequence $(1/n)$ lies in $\mathbb{R} - \{0\}$ and tends to 0, but

$$f(1/n) = \frac{1}{1/n} = n \to \infty \quad \text{as } n \to \infty.$$

Hence $f(x) = 1/x$ does not tend to a limit as x tends to 0.

Here are some limits of functions for you to consider. Recall that $\lfloor x \rfloor$ is the integer part of x.

Exercise F1

Determine whether each of the following limits exists, and evaluate those limits which do exist.

(a) $\displaystyle\lim_{x\to 0} \frac{x^2 + x}{x}$ (b) $\displaystyle\lim_{x\to 1} \lfloor x \rfloor$ (c) $\displaystyle\lim_{x\to 0} \log |x|$

1.2 Limits and continuity

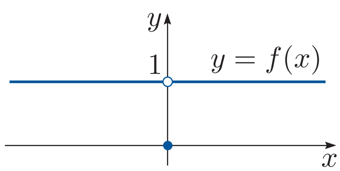


Figure 6 The graph of a function which takes the value 1 except at $x = 0$.

Consider the function

$$f(x) = \begin{cases} 1, & x \neq 0, \\ 0, & x = 0, \end{cases}$$

whose graph is shown in Figure 6. Does this function tend to a limit as x tends to 0 and, if so, what is the limit? Well, if (x_n) is any null sequence with non-zero terms, then

$$f(x_n) = 1, \quad \text{for } n = 1, 2, \ldots,$$

so

$$f(x_n) \to 1 \quad \text{as } n \to \infty.$$

Hence

$$\lim_{x\to 0} f(x) = 1.$$

This example illustrates the fact that the value of a limit $\lim_{x\to c} f(x)$ is not affected by the value of $f(c)$, if f happens to be defined at c.

However, the following theorem shows that if f is defined and *continuous* at c, then the value of the limit must be $f(c)$, and the converse statement is also true.

Theorem F2

Let f be a function defined on an open interval I, with $c \in I$. Then

f is continuous at c

if and only if

$$\lim_{x \to c} f(x) = f(c).$$

The proof of Theorem F2 uses the fact that, in this situation, the definition of continuity of f at c, from Subsection 2.1 of Unit D4, is almost identical to the definition of the existence of $\lim\limits_{x \to c} f(x)$, with this limit equal to $f(c)$. The only difference is that, in the former case, we allow the terms of the sequences (x_n) which appear in the definition to equal c. We omit the details of this proof.

Theorem F2 makes it easy to calculate many limits of continuous functions. For example, to determine

$$\lim_{x \to 2}(3x^5 - 5x^2 + 1),$$

we use the fact that the function $f(x) = 3x^5 - 5x^2 + 1$ is continuous on $\mathbb{R}$, since f is a polynomial. Hence, by Theorem F2,

$$\lim_{x \to 2}(3x^5 - 5x^2 + 1) = f(2) = 77.$$

As a further example, you saw in Worked Exercise D48 in Unit D4 that the function

$$f(x) = \begin{cases} x^2 \sin(1/x), & x \neq 0, \\ 0, & x = 0, \end{cases}$$

is continuous at 0. Thus, by Theorem F2,

$$\lim_{x \to 0} x^2 \sin(1/x) = 0.$$

On the other hand, we saw in Worked Exercise F2(a) that

$$\lim_{x \to 0} \sin(1/x) \text{ does not exist.}$$

It follows from Theorem F2 that, no matter how we try to extend the domain of the function $f(x) = \sin(1/x)$ to include $x = 0$, we can never obtain a continuous function.

Exercise F2

Use Theorem F2 to determine the following limits.

(a) $\lim\limits_{x \to 2} \sqrt{x}$ (b) $\lim\limits_{x \to \pi/2} \sqrt{\sin x}$ (c) $\lim\limits_{x \to 1} \dfrac{e^x}{1 + x}$

In the remainder of this unit we use Theorem F2 often, but we do not always refer to it explicitly.

1.3 Rules for limits

As you might expect from your experience with sequences, series and continuous functions, limits of functions can often be found by using various rules. First we state the Combination Rules for limits. These can be deduced from the corresponding rules for sequences; we omit the details.

Theorem F3 Combination Rules for limits

If $\lim_{x\to c} f(x) = l$ and $\lim_{x\to c} g(x) = m$, then:

Sum Rule $\lim_{x\to c}(f(x) + g(x)) = l + m$

Multiple Rule $\lim_{x\to c} \lambda f(x) = \lambda l$, for $\lambda \in \mathbb{R}$

Product Rule $\lim_{x\to c} f(x)g(x) = lm$

Quotient Rule $\lim_{x\to c} f(x)/g(x) = l/m$, provided that $m \neq 0$.

For example, since

$$\lim_{x\to 0} \frac{\sin x}{x} = 1 \quad \text{and} \quad \lim_{x\to 0}(x^2 + 1) = 1,$$

we have, by the Combination Rules,

$$\lim_{x\to 0}\left(\frac{\sin x}{x} + 2(x^2+1)\right) = 1 + 2\times 1 = 3.$$

Next we discuss the composition of limits. For example, consider the behaviour of

$$\frac{\sin(x^2)}{x^2},$$

as x tends to 0. This function can be written in the form

$$\frac{\sin u}{u}, \quad \text{where } u = x^2.$$

Now

$$u = x^2 \to 0 \text{ as } x \to 0 \quad \text{and} \quad \frac{\sin u}{u} \to 1 \text{ as } u \to 0,$$

which suggests that

$$\frac{\sin(x^2)}{x^2} \to 1 \text{ as } x \to 0.$$

To justify this composition of limits, we use the following Composition Rule. This can be proved using properties of convergent sequences, limits and continuous functions that you have already met, but we do not give the details here.

Theorem F4 Composition Rule for limits

If $\lim_{x\to c} f(x) = l$ and $\lim_{x\to l} g(x) = L$, then

$$\lim_{x\to c} g(f(x)) = L,$$

provided that

either $f(x) \neq l$, for all x in some $N_r(c)$, where $r > 0$,

or g is defined at l and continuous at l.

Remarks

1. When using this rule, it is important to remember that the limit for g is as $x \to l$ and not as $x \to c$. This is perhaps easier to see if we rewrite the theorem as follows:

   If

   $$f(x) \to l \text{ as } x \to c$$

   and

   $$g(x) \to L \text{ as } x \to l,$$

   then

   $$g(f(x)) \to L \text{ as } x \to c.$$

2. Before you use the Composition Rule, you should check that one of the two provisos to Theorem F4 holds. (Sometimes both provisos will hold, but you only ever need to check that one does.)

Perhaps surprisingly, the Composition Rule is *false* if we omit both of the provisos. For example, if

$$f(x) = 1 \quad \text{and} \quad g(x) = \begin{cases} 2, & x \neq 1, \\ 0, & x = 1, \end{cases}$$

then

$$\lim_{x\to 0} f(x) = 1 \quad \text{and} \quad \lim_{x\to 1} g(x) = 2$$

so the Composition Rule would give $\lim_{x\to 0} g(f(x)) = 2$. However, because $f(x) = 1$ for all values of x and $g(1) = 0$, we actually have

$$\lim_{x\to 0} g(f(x)) = 0 \neq 2.$$

For this example, neither of the provisos to Theorem F4 holds, because there is no punctured neighbourhood of 0 in which $f(x) \neq 1$, and g is not continuous at 1.

However, in nearly all cases in practice and in all the examples you will meet in this module, at least one of the provisos holds (though you should check this).

This leads to the following strategy for using the Composition Rule.

Strategy F2

To use the Composition Rule to evaluate a limit of a function of the form $g(f(x))$ as $x \to c$, do the following.

1. Substitute $u = f(x)$ and show that, for some l,
$$u = f(x) \to l \text{ as } x \to c.$$
2. Show that, for some L,
$$g(u) \to L \text{ as } u \to l.$$
3. Check that one of the provisos holds.
4. Deduce that
$$g(f(x)) \to L \text{ as } x \to c.$$

The following worked exercise illustrates how to apply this strategy.

Worked Exercise F3

Determine the following limits.

(a) $\displaystyle\lim_{x\to 0} \frac{\sin\left(\frac{1}{2}x\right)}{\frac{1}{2}x}$ (b) $\displaystyle\lim_{x\to 0} \left(1 + \left(\frac{\sin x}{x}\right)^2\right)$

Solution

(a) We want to use Strategy F2 so we identify functions f and g such that $\dfrac{\sin\left(\frac{1}{2}x\right)}{\frac{1}{2}x} = g(f(x))$.

We can write
$$\frac{\sin\left(\frac{1}{2}x\right)}{\frac{1}{2}x} = g(f(x)), \quad \text{where } f(x) = \tfrac{1}{2}x \text{ and } g(x) = \frac{\sin x}{x}.$$

We now follow the steps of Strategy F2.

Substituting $u = f(x) = \frac{1}{2}x$, we have
$$u = \tfrac{1}{2}x \to u(0) = 0 \text{ as } x \to 0,$$
since u is continuous at 0, and
$$g(u) = \frac{\sin u}{u} \to 1 \text{ as } u \to 0.$$
The first proviso of the Composition Rule holds because $f(x) \neq 0$ for $x \in N_1(0)$, for example.

Notice that the second proviso does not hold in this case, since g is undefined at 0. However, we only need one of the provisos to hold to use the Composition Rule.

Thus, by the Composition Rule,

$$g(f(x)) = \frac{\sin\left(\frac{1}{2}x\right)}{\frac{1}{2}x} \to 1 \text{ as } x \to 0.$$

(b) We can write

$$1 + \left(\frac{\sin x}{x}\right)^2 = g(f(x)),$$

where

$$f(x) = \frac{\sin x}{x} \text{ and } g(x) = 1 + x^2.$$

Substituting $u = f(x) = \dfrac{\sin x}{x}$, we obtain

$$u = \frac{\sin x}{x} \to 1 \text{ as } x \to 0,$$
$$g(u) = 1 + u^2 \to 1 + 1 = 2 \text{ as } u \to 1,$$

since g is continuous at 1, which also tells us that the second proviso to the Composition Rule holds.

Alternatively you could have noted that the first proviso holds, since $f(x) \neq 1$ for $x \in N_1(0)$, for example.

Thus, by the Composition Rule,

$$g(f(x)) = 1 + \left(\frac{\sin x}{x}\right)^2 \to 2 \text{ as } x \to 0.$$

Exercise F3

Use the Combination Rules and the Composition Rule to determine the following limits.

(a) $\displaystyle\lim_{x\to 0} \frac{\sin x}{2x + x^2}$ (b) $\displaystyle\lim_{x\to 0} \frac{\sin(\sin x)}{\sin x}$ (c) $\displaystyle\lim_{x\to 0} \left(\frac{x}{\sin x}\right)^{1/2}$

(d) $\displaystyle\lim_{x\to 0} \frac{1 - \cos x}{x}$

Hint: In part (d), use the identity $\cos x = 1 - 2\sin^2\left(\frac{1}{2}x\right)$.

There is also a Squeeze Rule for limits, analogous to the Squeeze Rules for sequences and continuous functions, whose proof we omit. This is illustrated in Figure 7.

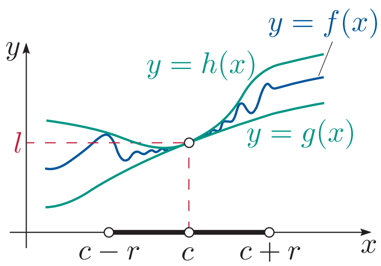


Figure 7 The Squeeze Rule for limits

Theorem F5 Squeeze Rule for limits

Let f, g and h be functions defined on $N_r(c)$, for some $r > 0$. If

(a) $g(x) \leq f(x) \leq h(x)$, for $x \in N_r(c)$

(b) $\lim\limits_{x\to c} g(x) = \lim\limits_{x\to c} h(x) = l$,

then

$$\lim_{x\to c} f(x) = l.$$

In the proof of Theorem F1, we showed that $\lim\limits_{x\to 0} (\sin x)/x = 1$, using the inequalities

$$\cos x \leq \frac{\sin x}{x} \leq 1, \quad \text{for } 0 < |x| < \pi/2.$$

This was, in essence, an application of the Squeeze Rule for limits, with $f(x) = (\sin x)/x$, $g(x) = \cos x$ and $h(x) = 1$. Since $\lim\limits_{x\to 0} \cos x = 1$, the result follows.

In the next exercise the Squeeze Rule is used to establish another important limit.

Exercise F4

(a) Use the inequalities

$$1 + x \leq e^x \leq \frac{1}{1-x}, \quad \text{for } |x| < 1,$$

(proved in Corollary D49 in Unit D4) to show that

$$1 - \frac{|x|}{1-x} \leq \frac{e^x - 1}{x} \leq 1 + \frac{|x|}{1-x}, \quad \text{for } 0 < |x| < 1.$$

(b) Deduce from part (a) that

$$\lim_{x\to 0} \frac{e^x - 1}{x} = 1.$$

The limit found in Exercise F4 is one of the basic limits we often use. Here we record three such limits for future reference. The first was proved in Theorem F1, the second in Exercise F3(d) and the third in Exercise F4(b).

Theorem F6 Three basic limits

(a) $\displaystyle\lim_{x\to 0} \frac{\sin x}{x} = 1$

(b) $\displaystyle\lim_{x\to 0} \frac{1-\cos x}{x} = 0$

(c) $\displaystyle\lim_{x\to 0} \frac{e^x - 1}{x} = 1$

1.4 One-sided limits

Earlier we mentioned that $\lim_{x\to 0} \sqrt{x}$ is not defined because the function $f(x) = \sqrt{x}$ is not defined on any punctured neighbourhood of 0. However, this function does tend to 0 as x tends to 0 *from the right*.

Definitions

Let f be a function defined on $(c, c+r)$, for some $r > 0$. Then $f(x)$ **tends to the limit l as x tends to c from the right** if

for each sequence (x_n) in $(c, c+r)$ such that $x_n \to c$,
$f(x_n) \to l$.

In this case, we write

$$\lim_{x\to c^+} f(x) = l \quad \text{or} \quad f(x) \to l \text{ as } x \to c^+.$$

There is a similar definition for a limit **as x tends to c from the left**, in which $(c, c+r)$ is replaced by $(c-r, c)$. In this case, we write

$$\lim_{x\to c^-} f(x) = l \quad \text{or} \quad f(x) \to l \text{ as } x \to c^-.$$

We also refer to

$$\lim_{x\to c^+} f(x) \quad \text{and} \quad \lim_{x\to c^-} f(x)$$

as **right** and **left limits**, respectively.

Sometimes both right and left limits exist but are different, as you will see in the next worked exercise.

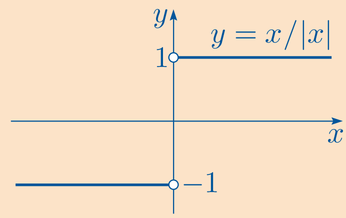
Worked Exercise F4

Prove that the following function tends to different limits as x tends to 0 from the right and from the left.

$$f(x) = \frac{x}{|x|} \quad (x \in \mathbb{R} - \{0\})$$

Solution

The graph of the function is shown below.



The function f is defined on $(0, 1)$ and $f(x) = 1$ on this open interval.

We could take any open interval of the form $(0, r)$ with $r > 0$ here.

Thus if (x_n) is a null sequence in $(0, 1)$, then

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} 1 = 1.$$

Hence $\displaystyle\lim_{x\to 0^+} f(x) = 1$.

Similarly, f is defined on $(-1, 0)$ and $f(x) = -1$ on this interval.

We could take any open interval of the form $(-r, 0)$ with $r > 0$ here.

Thus if (x_n) is a null sequence in $(-1, 0)$, then

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} -1 = -1.$$

Hence $\displaystyle\lim_{x\to 0^-} f(x) = -1$.

Since $-1 \neq 1$, the limits of $f(x)$ as x tends to 0 from the right and from the left are different.

The relationship between one-sided limits and ordinary limits is given by the following result, whose proof we omit.

Theorem F7

Let the function f be defined on $N_r(c)$, for some $r > 0$. Then

$$\lim_{x \to c} f(x) = l$$

if and only if

$$\lim_{x \to c^+} f(x) = \lim_{x \to c^-} f(x) = l.$$

Analogues of the Combination Rules, the Composition Rule (and Strategy F2) and the Squeeze Rule can also be used to determine one-sided limits. In the statements of these rules, we simply replace $\lim\limits_{x \to c}$ by $\lim\limits_{x \to c^+}$ or $\lim\limits_{x \to c^-}$, and replace $N_r(c)$ by $(c, c + r)$ or $(c - r, c)$, as appropriate. Also, Strategy F1 can be adapted to show that a one-sided limit does *not* exist; the sequences (x_n) and (y_n) must be chosen to tend to c from the right or from the left, as appropriate.

There is also a version of Theorem F2 for one-sided limits, as follows.

Theorem F8

Let f be a function whose domain is an interval I with a finite left-hand endpoint c that lies in I. Then

f is continuous at c

if and only if

$$\lim_{x \to c^+} f(x) = f(c).$$

Remarks

1. In Theorem F8, the interval I can have any of the forms $[c, \infty)$, $[c, b)$ or $[c, b]$, where $b > c$.
2. There is an analogous result to Theorem F8 for left limits.

It follows from Theorem F8 that $\lim\limits_{x \to 0^+} \sqrt{x} = 0$, as claimed earlier, since $f(x) = \sqrt{x}$ has domain $[0, \infty)$ and is continuous at 0.

Exercise F5

Prove the following.

(a) $\displaystyle \lim_{x \to 0^+} \left(\frac{\sin x}{x} + \sqrt{x} \right) = 1$ (b) $\displaystyle \lim_{x \to 0^+} \frac{\sin \sqrt{x}}{\sqrt{x}} = 1$

2 Asymptotic behaviour of functions

In our discussion of graph sketching in Unit A4 *Real functions, graphs and conics*, we described several types of *asymptotic behaviour* (that is, behaviour of a function when the domain variable or codomain variable becomes arbitrarily large), such as:

$$\frac{1}{x} \to \infty \text{ as } x \to 0^+ \quad \text{and} \quad e^x \to \infty \text{ as } x \to \infty.$$

In this section we define such statements formally and describe various relationships between them.

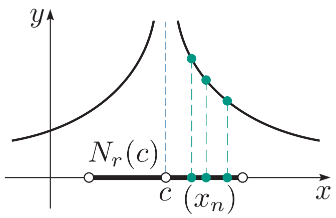


Figure 8 A function for which $f(x)$ tends to infinity as x tends to c

2.1 Functions which tend to infinity

In Section 1 we defined $f(x) \to l$ as $x \to c$ for a finite limit l in terms of the behaviour of sequences. We can define

$$f(x) \to \infty \text{ as } x \to c$$

in a similar way. This is illustrated in Figure 8.

Definition

Let the function f be defined on $N_r(c)$, for some $r > 0$. Then $f(x)$ **tends to ∞ as x tends to c** if

for each sequence (x_n) in $N_r(c)$ such that $x_n \to c$,
$f(x_n) \to \infty$.

In this case, we write

$$f(x) \to \infty \text{ as } x \to c.$$

Remarks

1. In this module we do not use the notation $\lim\limits_{x \to c} f(x) = \infty$ as this can give the misleading impression that infinity can be treated in the same way as a finite limit. Algebraic manipulations of expressions involving ∞ are a common error in false proofs – as you saw, for example, for a series that is not convergent in Worked Exercise D29 in Unit D3 *Series*.
2. The statements

   $$f(x) \to -\infty \text{ as } x \to c,$$
   $$f(x) \to \infty \text{ (or } -\infty) \text{ as } x \to c^+ \text{ (or } c^-),$$

   are defined similarly, with ∞ replaced by $-\infty$ and $N_r(c)$ replaced by the open interval $(c, c + r)$ or $(c - r, c)$, where $r > 0$, as appropriate.

There is a version of the Reciprocal Rule which relates functions that tend to infinity and functions that tend to 0. (You met the Reciprocal Rule for sequences in Subsection 4.3 of Unit D2.)

Theorem F9 Reciprocal Rule for limits

If the function f satisfies the conditions

1. $f(x) > 0$ for $x \in N_r(c)$, for some $r > 0$
2. $f(x) \to 0$ as $x \to c$,

then

$$1/f(x) \to \infty \quad \text{as } x \to c.$$

For example,

$$1/x^2 \to \infty \quad \text{as } x \to 0,$$

because $f(x) = x^2 > 0$ for $x \in \mathbb{R} - \{0\}$, and $\lim\limits_{x \to 0} x^2 = 0$; see Figure 9.

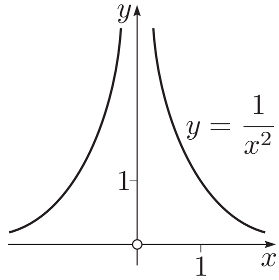


Figure 9 The graph of $y = 1/x^2$

The Reciprocal Rule can also be applied with $x \to c$ replaced by $x \to c^+$ or $x \to c^-$, and $N_r(c)$ replaced by $(c, c+r)$ or $(c-r, c)$, as appropriate. For example, we have

$$1/x \to \infty \quad \text{as } x \to 0^+,$$

because $f(x) = x > 0$ for $x \in (0, \infty)$, and $\lim\limits_{x \to 0^+} x = 0$; see Figure 10.

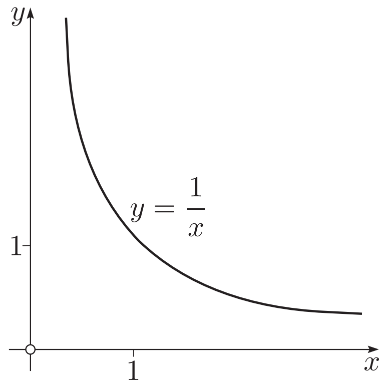


Figure 10 The graph of $y = 1/x$

Exercise F6

Prove that

(a) $\dfrac{1}{|x|} \to \infty$ as $x \to 0$ (b) $\dfrac{\sin x}{x^3} \to \infty$ as $x \to 0$

(c) $\dfrac{1}{x^3 - 1} \to \infty$ as $x \to 1^+$.

There are also versions of the Combination Rules and the Squeeze Rule for functions which tend to ∞ (or $-\infty$) as x tends to c, c^+ or c^-. Here we state the Combination Rules for functions which tend to ∞ as x tends to c.

Theorem F10 Combination Rules for functions which tend to infinity

If $f(x) \to \infty$ as $x \to c$ and $g(x) \to \infty$ as $x \to c$, then:

Sum Rule $f(x) + g(x) \to \infty$ as $x \to c$

Multiple Rule $\lambda f(x) \to \infty$ as $x \to c$, for $\lambda \in \mathbb{R}^+$

Product Rule $f(x)g(x) \to \infty$ as $x \to c$.

These rules are analogous to the corresponding rules for sequences which tend to infinity; see Subsection 4.3 of Unit D2.

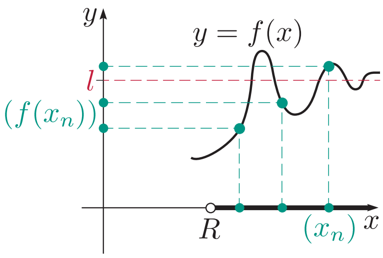


Figure 11 A function for which $f(x) \to l$ as $x \to \infty$

2.2 Behaviour as x tends to infinity

Next, we define various types of behaviour of real functions $f(x)$ as $x \to \infty$ or as $x \to -\infty$. To avoid repetition, in the following definition we allow the letter l to denote either a real number or one of the symbols ∞ or $-\infty$. The definition is illustrated in Figure 11.

Definition

Let the function f be defined on (R, ∞), for some real number R. Then $f(x)$ **tends to l as x tends to ∞** if

for each sequence (x_n) in (R, ∞) such that $x_n \to \infty$,
$f(x_n) \to l$.

In this case, we write

$$f(x) \to l \quad \text{as } x \to \infty.$$

The statement

$$f(x) \to l \quad \text{as } x \to -\infty$$

is defined similarly, with ∞ replaced by $-\infty$, and (R, ∞) replaced by $(-\infty, R)$. Note that

$$f(x) \to l \quad \text{as } x \to -\infty$$

is equivalent to

$$f(-x) \to l \quad \text{as } x \to \infty.$$

When l is a real number, we also use the notations

$$\lim_{x \to \infty} f(x) = l \quad \text{and} \quad \lim_{x \to -\infty} f(x) = l.$$

Once again, we can use versions of the Reciprocal Rule and the Combination Rules to obtain results about the behaviour of given functions as $x \to \infty$ or $-\infty$. The new versions of these rules are obtained from the Reciprocal Rule and the Combination Rules in Subsection 2.1 and the Combination Rules in Subsection 1.3 by replacing c by ∞ or $-\infty$, and $N_r(c)$ by (R, ∞) or $(-\infty, R)$, as appropriate.

Many results about the behaviour of functions $f(x)$ as x tends to ∞ or $-\infty$ are derived from the following two basic facts, often by using the Combination Rules and the Reciprocal Rule.

Theorem F11 Basic asymptotic behaviour

If $n \in \mathbb{N}$, then

(a) $x^n \to \infty$ as $x \to \infty$

(b) $\dfrac{1}{x^n} \to 0$ as $x \to \infty$.

We can use Theorem F11, together with the Combination Rules and the Reciprocal Rule, to determine the asymptotic behaviour of various functions defined by quotients. This is similar to determining the behaviour of sequences defined by quotients (see Subsection 3.2 of Unit D2), and we give a corresponding definition of the *dominant term* of a quotient that suits the present context.

Definition

The **dominant term** of a quotient involving the real variable x is the term in x (without its coefficient) which eventually has the largest absolute value.

For example, consider the behaviour of $x/(x^2+1)$ as $x \to \infty$. Here the dominant term is x^2, so we divide both the numerator and the denominator by x^2 to give

$$\frac{x}{x^2+1} = \frac{1/x}{1+1/x^2} \to \frac{0}{1+0} = 0 \quad \text{as } x \to \infty,$$

by Theorem F11(b) and the Combination Rules.

Exercise F7

Prove that:

(a) $\displaystyle\lim_{x\to\infty} \frac{2x^3+x}{x^3} = 2$ (b) $\dfrac{2x^3+1}{x^2} \to \infty$ as $x \to \infty$.

There are also versions of the Squeeze Rule for functions as x tends to infinity, which have some important applications. You met the corresponding versions of the Squeeze Rule for sequences in Theorems D10 and D18 in Sections 3 and 4 of Unit D2.

Theorem F12 Squeeze Rule for functions as $x \to \infty$

Let f, g and h be functions defined on some interval (R, ∞).

(a) If f, g and h satisfy the conditions

1. $g(x) \leq f(x) \leq h(x)$, for $x \in (R, \infty)$
2. $\lim_{x\to\infty} g(x) = \lim_{x\to\infty} h(x) = l$

where l is a real number, then

$$\lim_{x\to\infty} f(x) = l.$$

(b) If f and g satisfy the conditions

1. $f(x) \geq g(x)$, for $x \in (R, \infty)$
2. $g(x) \to \infty$ as $x \to \infty$

then

$$f(x) \to \infty \quad \text{as } x \to \infty.$$

Exercise F8

Use the Squeeze Rule to determine the behaviour of the following function as $x \to \infty$.

$$f(x) = \frac{\sin(1/x)}{x}$$

Hint: Use the fact that $-1 \leq \sin(1/x) \leq 1$ for $x \neq 0$.

In Subsection 1.3 we gave the Composition Rule for limits and Strategy F2 for using it. We can also use this strategy to deduce the asymptotic behaviour of composites of functions which have any of the types of asymptotic behaviour introduced in this unit, provided that we allow the letters l and L to denote either a real number or one of the symbols ∞ or $-\infty$. Notice that if l is either ∞ or $-\infty$, then the first proviso of the Composition Rule is automatically satisfied, so there is no need to check this.

For example, consider the asymptotic behaviour of $\dfrac{\sin(1/x^2)}{x^2}$ as $x \to \infty$. We can write

$$\frac{\sin(1/x^2)}{x^2} = g(f(x)),$$

where $f(x) = x^2$ and $g(x) = \dfrac{\sin(1/x)}{x}$. Substituting $u = f(x) = x^2$, we have

$$u = f(x) = x^2 \to \infty \quad \text{as } x \to \infty \quad \text{(by Theorem F11)},$$

$$g(u) = \frac{\sin(1/u)}{u} \to 0 \quad \text{as } u \to \infty \quad \text{(by Exercise F8)}.$$

Thus we deduce by Strategy F2 that

$$g(f(x)) = \frac{\sin(1/x^2)}{x^2} \to 0 \text{ as } x \to \infty.$$

In the next theorem, we collect together several standard results about the behaviour of particular functions as $x \to \infty$.

Theorem F13

(a) If $a_0, a_1, \dots, a_{n-1} \in \mathbb{R}$, where $n \in \mathbb{N}$, and

$$p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0,$$

then

$$p(x) \to \infty \text{ as } x \to \infty \quad \text{and} \quad \frac{1}{p(x)} \to 0 \text{ as } x \to \infty.$$

(b) For each $n = 0, 1, 2, \dots$, we have

$$\frac{e^x}{x^n} \to \infty \text{ as } x \to \infty \quad \text{and} \quad \frac{x^n}{e^x} \to 0 \text{ as } x \to \infty.$$

(c) We have

$$\log x \to \infty \text{ as } x \to \infty,$$

but, for each constant $a > 0$, we have

$$\frac{\log x}{x^a} \to 0 \text{ as } x \to \infty.$$

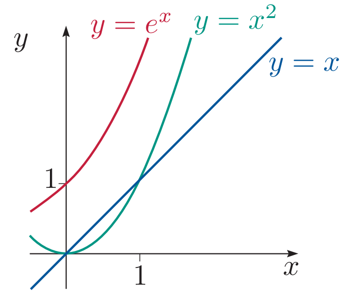


Figure 12 The graphs of $y = e^x$, $y = x^2$ and $y = x$

Part (b) tells us that, as x tends to infinity, e^x tends to infinity *faster* than any positive integer power of x, as illustrated in Figure 12. On the other hand, part (c) tells us that, as x tends to infinity, $\log x$ tends to infinity *more slowly* than any positive power of x. Thus, in part (b) the dominant term is e^x, whereas in part (c) it is x^a.

Proof of Theorem F13

(a) We use the fact that all the zeros of the polynomial p must lie in the interval $(-M, M)$, where $M = 1 + \max\{|a_{n-1}|, \dots, |a_1|, |a_0|\}$, and

$$p(x) > 0, \quad \text{for } x \in (M, \infty). \tag{5}$$

This was proved in Theorem D54 of Unit D4.

Now for $x \neq 0$,

$$p(x) = x^n \left(1 + \frac{a_{n-1}}{x} + \cdots + \frac{a_0}{x^n}\right).$$

By Theorem F11(b) and the Combination Rules,

$$1 + \frac{a_{n-1}}{x} + \cdots + \frac{a_0}{x^n} \to 1 + 0 + \cdots + 0 = 1 \text{ as } x \to \infty. \tag{6}$$

Thus, for $x \in (M, \infty)$, we have

$$\frac{1}{p(x)} = \frac{1/x^n}{1 + a_{n-1}/x + \cdots + a_0/x^n} \to \frac{0}{1} = 0 \quad \text{as } x \to \infty,$$

by statement (6), Theorem F11(b) and the Quotient Rule for limits. We deduce, by inequality (5) and the Reciprocal Rule, that

$$p(x) \to \infty \quad \text{as } x \to \infty.$$

(b) Let n be a fixed non-negative integer. We use the series representation

$$e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \frac{x^{n+1}}{(n+1)!} + \cdots,$$

which was shown to be valid for $x \geq 0$ in Subsection 4.1 of Unit D3. Since $x \geq 0$, all the terms in the above series are non-negative, so

$$e^x \geq \frac{x^{n+1}}{(n+1)!}, \qquad \text{for } x \geq 0.$$

Hence, for $x > 0$,

$$\frac{e^x}{x^n} \geq \frac{x}{(n+1)!} \quad \text{and} \quad 0 \leq \frac{x^n}{e^x} \leq \frac{(n+1)!}{x}.$$

It follows by Theorem F11, the Multiple Rule and the Squeeze Rule that

$$\frac{e^x}{x^n} \to \infty \quad \text{as } x \to \infty \quad \text{and} \quad \frac{x^n}{e^x} \to 0 \quad \text{as } x \to \infty.$$

(c) It was shown in Subsection 4.2 of Unit D4 that the function $x \longmapsto \log x$ is a strictly increasing inverse of the exponential function, with domain $(0, \infty)$ and range $\mathbb{R}$. We deduce that

$$\log x \to \infty \quad \text{as } x \to \infty.$$

Now let a be any positive constant. Since $x^a = \exp(a \log x)$, we make the substitution $t = a \log x$, so that $x^a = e^t$. For $x > 0$, this gives

$$\frac{\log x}{x^a} = \frac{t/a}{e^t} = \frac{t}{ae^t}. \tag{7}$$

Since $a > 0$, we have

$$t = a \log x \to \infty \quad \text{as } x \to \infty, \tag{8}$$

and, by part (b) with $n = 1$ and the Multiple Rule,

$$\frac{t}{ae^t} \to 0 \quad \text{as } t \to \infty. \tag{9}$$

Hence, by statements (7), (8) and (9), together with the Composition Rule, we have

$$\frac{\log x}{x^a} = \frac{t}{ae^t} \to 0 \quad \text{as } x \to \infty,$$

as required. ■

Exercise F9

Use the results of Theorem F13 and appropriate rules to determine the behaviour of the following functions as $x \to \infty$.

(a) $f(x) = \dfrac{e^x}{x^2} + \dfrac{3x^2}{\log x}$ (b) $f(x) = \dfrac{\log x}{e^x}$ (c) $f(x) = \dfrac{2e^x - x^2}{e^x + \log x}$

Hint: In part (b), express $(\log x)/e^x$ in terms of $(\log x)/x$ and x/e^x.

Exercise F10

Prove that:

(a) $e^{x^2}/x^2 \to \infty$ as $x \to \infty$ (b) $\log(\log x) \to \infty$ as $x \to \infty$

(c) $x \sin(1/x) \to 1$ as $x \to \infty$.

Hint: In part (c), use the substitution $u = 1/x$.

3 Continuity: the classical definition

In Book D you met a definition of continuity based on sequences, and saw that most familiar functions are continuous on their domains, a fact which is not at all surprising. However, there are many functions of interest for which it is more difficult to establish continuity (or discontinuity). Consider, for instance, the **Riemann function**, which has domain $\mathbb{R}$ and rule

$$f(x) = \begin{cases} 1/q, & \text{if } x \text{ is a rational } p/q, \text{ where } q > 0, \\ 0, & \text{if } x \text{ is irrational.} \end{cases}$$

Note that, in this section, we assume that all rationals p/q are expressed in lowest terms; that is, the greatest common factor of p and q is 1.

It follows from the definition that, for example, $f(2/3) = 1/3$ and $f(\sqrt{2}) = 0$. By plotting the values of $f(x)$ at different values of x we can produce a sketch of the graph of the Riemann function as shown in Figure 13.

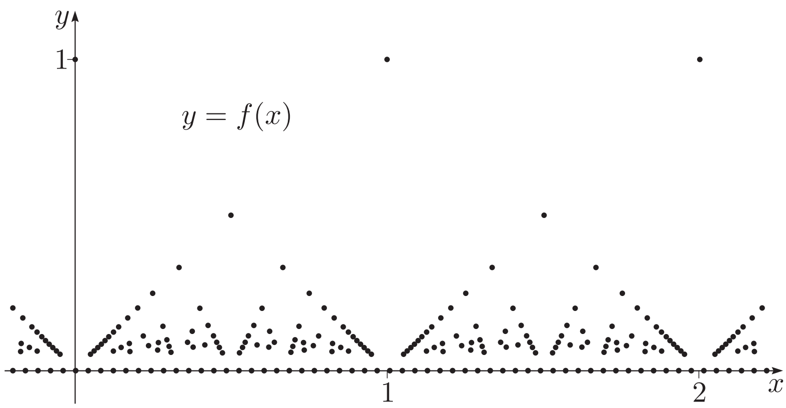


Figure 13 A sketch of the graph of the Riemann function

From this sketch of the graph of f it is not clear whether the Riemann function is continuous at any point of $\mathbb{R}$, but you will see (in Subsection 3.2) that in fact it is continuous at infinitely many points and discontinuous at infinitely many points! When dealing with such unusual functions, it is useful to have available the alternative definition of continuity which is introduced in Subsection 3.1. This definition looks more abstract but is more effective in some cases.

The emergence of rigorous analysis

The classical definition of continuity described in this section emerged towards the end of the nineteenth century after many years of debate amongst mathematicians about the rigorous formulation of analysis as the foundation of calculus. At this time, the informal approach used in the eighteenth century, for instance by Euler, was increasingly found to be inadequate. For example, around 1820, Joseph Fourier (1768–1830) used functions defined by infinite series to solve problems in the theory of heat. The properties of these series raised challenging questions about the meaning of convergence. This led to questions about the definitions of continuity, limits, differentiation and integration, and even the nature of the real numbers. These questions were not properly resolved until about 1870, after contributions by many mathematicians, including Bolzano, Cauchy, Riemann, Dirichlet, Dedekind, Weierstrass and Cantor.

3.1 The ε-δ definition of continuity

The sequential definition of continuity that you studied in Unit D4 states that the function $f : A \longrightarrow \mathbb{R}$ is continuous at c, where $c \in A$, if

for each sequence (x_n) in A such that $x_n \to c$,
$f(x_n) \to f(c)$.

This definition uses sequences to formalise the intuitive idea that $f(x)$ approaches $f(c)$ as x approaches the point c in any manner.

The new definition that we study in this unit formalises this idea in a somewhat different way, which we can describe in words as follows:

we can make $f(x)$ as close as we wish to $f(c)$ by ensuring that x is close enough to c.

The 'closeness' in this description is measured by two variables, ε in the codomain and δ in the domain, which represent 'small' positive numbers.

Definition

Let the function f have domain A and let $c \in A$. Then f is **continuous** at c if

$$\text{for each } \varepsilon > 0, \text{ there exists } \delta > 0 \text{ such that}$$
$$|f(x) - f(c)| < \varepsilon, \quad \text{for all } x \in A \text{ with } |x - c| < \delta. \tag{10}$$

Remarks

1. The above definition is quite subtle. It can be expressed in words as follows: no matter how small a positive number ε we are *given*, we can *choose* a positive number δ such that if the distance between x and c is less than δ, then the distance between $f(x)$ and $f(c)$ is less than ε. Thus statement (10) can be interpreted as an *implication*:

   $$\text{if } x \in A \text{ and } |x - c| < \delta, \text{ then } |f(x) - f(c)| < \varepsilon.$$

2. Note that

   $$\begin{array}{lll} |f(x) - f(c)| < \varepsilon & \text{is equivalent to} & f(c) - \varepsilon < f(x) < f(c) + \varepsilon, \\ |x - c| < \delta & \text{is equivalent to} & c - \delta < x < c + \delta, \end{array}$$

   as illustrated in Figure 14. This shows that as x gets closer to c, $f(x)$ gets closer to $f(c)$.

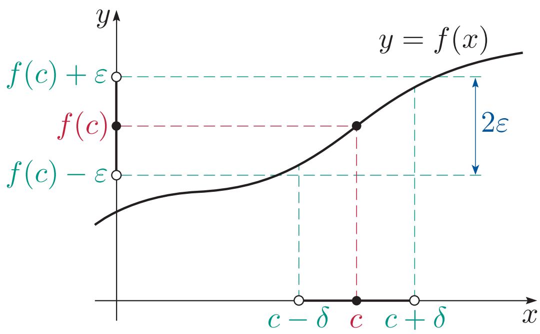


Figure 14 The ε-δ definition of continuity

3. Usually the value of δ that we choose in order to make statement (10) true depends on the given value of ε: the smaller ε is, the smaller δ has to be. The value of δ often depends also on the particular point c at which we are checking continuity.

We can interpret the task of finding a suitable choice of δ, when using this definition, as an 'ε-δ game' in which player A chooses a small positive number ε and then challenges player B to find a suitably small positive number δ for which statement (10) is true. (This is like the 'ε-N game' for null sequences that you met in Unit D2.) For example, suppose that $f(x) = x^2$ and $c = 0$. If player A chooses $\varepsilon = 1/4$ then player B can choose $\delta = 1/2$ (or any smaller value) because if

$$|x - 0| = |x| < 1/2,$$

then

$$|f(x) - f(0)| = |x^2| < 1/4.$$

This is illustrated in Figure 15.

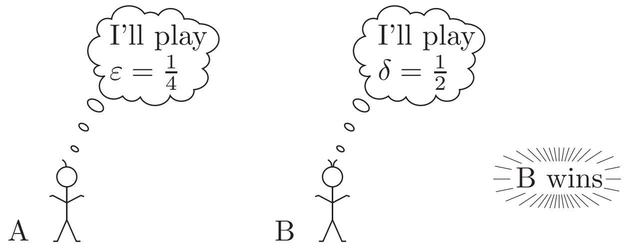


Figure 15 The ε-δ game for $f(x) = x^2$ at $c = 0$

In practice, we do not usually choose specific values for ε when proving continuity, but this game illustrates the ideas involved.

The method of applying the ε-δ definition of continuity depends on the nature of the function f. As an illustration, we apply the definition to polynomial functions, using the following strategy.

Strategy F3

To use the ε-δ definition to prove that a polynomial function f with domain A is continuous at a point $c \in A$, let $\varepsilon > 0$ be given and carry out the following steps.

1. Use algebraic manipulation to express the difference $f(x) - f(c)$ as a product of the form $(x - c)g(x)$.
2. Obtain an upper bound of the form $|g(x)| \leq M$, for $|x - c| \leq r$, where $r > 0$ is chosen so that $[c - r, c + r] \subset A$. (The Triangle Inequality is often useful here.)
3. Use the fact that $|f(x) - f(c)| \leq M|x - c|$, for $|x - c| \leq r$, to choose $\delta > 0$ such that

$$|f(x) - f(c)| < \varepsilon, \quad \text{for all } x \in A \text{ with } |x - c| < \delta.$$

The following worked exercise shows how Strategy F3 can be used.

Worked Exercise F5

Use the ε-δ definition to prove that $f(x) = x^2$ is continuous at $c = 2$.

Solution

The domain of f is $\mathbb{R}$.

Let $\varepsilon > 0$ be given. We want to choose $\delta > 0$, in terms of ε, such that

$$|f(x) - f(2)| < \varepsilon, \quad \text{for all } x \text{ with } |x - 2| < \delta. \qquad (*)$$

We follow the steps in Strategy F3.

1. First we write

$$f(x) - f(2) = x^2 - 4 = (x - 2)(x + 2).$$

   Writing f in this form will help us to show that if $|x - 2|$ is small, then $|f(x) - f(2)|$ is small.

2. Next we obtain an upper bound for $|x + 2|$ when x is near 2.

   We consider points for which $|x - 2| \leq r$ with $r = 1$; any $r > 0$ is suitable, but the resulting bounds will depend on r.

   If $|x - 2| \leq 1$, then x lies in the closed interval $[1, 3]$, so

$$\begin{aligned} |x + 2| &\leq |x| + 2 \qquad \text{(by the Triangle Inequality)} \\ &\leq 3 + 2 = 5. \end{aligned}$$

3. Hence

$$|f(x) - f(2)| \leq 5|x - 2|, \quad \text{for } |x - 2| \leq 1.$$

   So if $|x - 2| < \delta$, where $0 < \delta \leq 1$, then

$$|f(x) - f(2)| < 5\delta.$$

   Now $5\delta \leq \varepsilon$ if and only if $\delta \leq \frac{1}{5}\varepsilon$.

   We now choose δ so that *both* of the inequalities $0 < \delta \leq 1$ and $0 < \delta \leq \frac{1}{5}\varepsilon$ are satisfied.

   Thus, if we choose $\delta = \min\left\{1, \frac{1}{5}\varepsilon\right\}$, then

$$|f(x) - f(2)| < 5\delta \leq 5 \times \tfrac{1}{5}\varepsilon = \varepsilon, \quad \text{for all } x \text{ with } |x - 2| < \delta,$$

   which proves statement $(*)$.

Thus f is continuous at the point 2.

Exercise F11

Use the ε-δ definition to prove that $f(x) = x^3$ is continuous at $c = 1$.

Hint: Note that $x^3 - 1 = (x - 1)(x^2 + x + 1)$.

Next we verify that the two definitions of continuity are equivalent.

Theorem F14

The ε-δ definition and the sequential definition of continuity are equivalent.

Proof Let the function f have domain A, with $c \in A$. First we assume that f is continuous at c according to the ε-δ definition. We want to deduce that,

for each sequence (x_n) in A such that $x_n \to c$,

$$f(x_n) \to f(c). \tag{11}$$

Let $\varepsilon > 0$ be given. Then, by the ε-δ definition of continuity, there exists $\delta > 0$ such that

$$|f(x) - f(c)| < \varepsilon, \quad \text{for all } x \in A \text{ with } |x - c| < \delta. \tag{12}$$

Since $x_n \to c$, there exists an integer N such that

$$|x_n - c| < \delta, \quad \text{for all } n > N.$$

Hence, by statement (12),

$$|f(x_n) - f(c)| < \varepsilon, \quad \text{for all } n > N.$$

Thus statement (11) does indeed hold, so the sequential definition follows from the ε-δ definition.

Next suppose that f is continuous at c according to the sequential definition. We want to deduce that if $\varepsilon > 0$ is given, then there exists $\delta > 0$ such that statement (12) holds.

We use a proof by contradiction.

Suppose that, for some $\varepsilon > 0$, there is *no* such $\delta > 0$. Then statement (12) must be false with $\delta = 1$, $\delta = \frac{1}{2}$, $\delta = \frac{1}{3}$, and so on. Hence, for each $n \in \mathbb{N}$, there exists $x_n \in A$ with $|x_n - c| < 1/n$ such that

$$|f(x_n) - f(c)| \geq \varepsilon. \tag{13}$$

Now, the sequence (x_n) lies in A and $x_n \to c$. Thus, by the sequential definition of continuity, we have $\lim_{n\to\infty} f(x_n) = f(c)$, which contradicts inequality (13). We deduce that the ε-δ definition of continuity follows from the sequential definition. ■

3.2 Continuity of some unusual functions

It is natural to ask which is the 'better' definition of continuity. It is difficult to give a definitive answer, but on the whole:

- when proving the continuity of simpler functions the sequential definition is usually easier, whereas the ε-δ definition can work better with more complicated functions
- when proving discontinuity the sequential definition is usually easier.

For most of the functions you have met so far in this module, the points where the functions are continuous have been 'obvious'. However, there are many functions for which it is far less clear where they are continuous, if anywhere. In this subsection you will meet several interesting but quite complicated functions, and you will see that the ε-δ definition is an effective means of proving continuity, even when it is not possible to use Strategy F3.

The proofs in this subsection will take some effort to understand fully, so you may prefer to skim through them on a first reading and return to them as time permits. Do not be discouraged if you find them rather hard at first: reading proofs in analysis gets easier as you become more familiar with the sorts of arguments used.

The Dirichlet function and the Riemann function

The first function we consider has a simple definition, but is highly discontinuous. The **Dirichlet function** has domain $\mathbb{R}$ and rule

$$f(x) = \begin{cases} 1, & \text{if } x \text{ is rational,} \\ 0, & \text{if } x \text{ is irrational.} \end{cases}$$

The graph of f in Figure 16 looks rather like two parallel lines, but each line has infinitely many 'holes' in it!

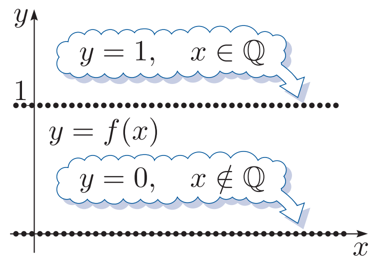


Figure 16 The Dirichlet function

Theorem F15

The Dirichlet function is discontinuous at every point of $\mathbb{R}$.

Proof Let c be any point of $\mathbb{R}$. We show that f is discontinuous at c by using the sequential definition of continuity.

We do this by showing that there are sequences that tend to c whose images under f tend to different limits. Recall the Density Property of $\mathbb{R}$ from Subsection 1.4 of Unit D1 *Numbers*: between any two real numbers, we can find both a rational number and an irrational number.

By the Density Property of $\mathbb{R}$, each open interval of the form

$$(c - 1/n, c + 1/n), \quad \text{where } n \in \mathbb{N},$$

contains a rational x_n and an irrational y_n. Considering the sequences (x_n) and (y_n), we have $x_n \to c$ and $y_n \to c$ by the Squeeze Rule for sequences, but

$$f(x_n) = 1 \quad \text{and} \quad f(y_n) = 0, \quad \text{for } n = 1, 2, \ldots.$$

Since $(f(x_n))$ and $(f(y_n))$ have different limits, f is discontinuous at c. ■

Our next function shows even stranger behaviour. The Riemann function, which we introduced at the start of this section, has domain $\mathbb{R}$ and rule

$$f(x) = \begin{cases} 1/q, & \text{if } x \text{ is a rational } p/q, \text{ where } q > 0, \\ 0, & \text{if } x \text{ is irrational.} \end{cases}$$

(Recall that, in this section, p/q is always expressed in its lowest terms.)

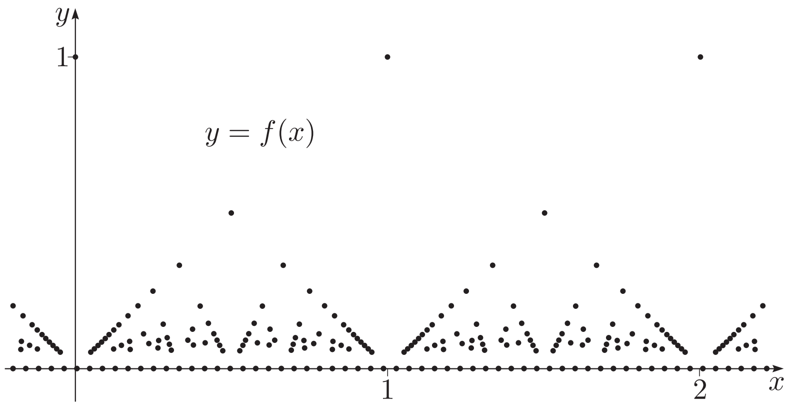


Figure 17 A sketch of the Riemann function

As mentioned earlier, it is not clear from the sketch of the graph of f in Figure 17 whether the Riemann function is continuous at any point of $\mathbb{R}$. In fact, it has the remarkable property that each open interval of $\mathbb{R}$ contains infinitely many points where f is continuous and infinitely many points where f is discontinuous. As you study the proof, notice how it illustrates the strengths of the two definitions of continuity.

Theorem F16

The Riemann function is discontinuous at each rational point of $\mathbb{R}$ and continuous at each irrational point.

Proof Here we prove discontinuity using the sequential definition and we prove continuity using the ε-δ definition.

First we prove that f is discontinuous at rational points.

Here we use Strategy D14 from Unit D4; that is, for each rational point c, we find a sequence that converges to c but whose images under f do not converge to $f(c)$.

Let $c = p/q$, with $q > 0$ (where p/q is expressed in lowest terms). Then, by the Density Property of $\mathbb{R}$, each open interval of the form

$$(c - 1/n, c + 1/n), \quad \text{where } n \in \mathbb{N},$$

contains an irrational number x_n. Considering the sequence (x_n), we have $x_n \to c$ and $f(x_n) = 0$, for $n = 1, 2, \ldots$. Since $f(c) = 1/q \neq 0$, we have $f(x_n) \nrightarrow f(c)$, so f is discontinuous at c.

Recall that the notation $\nrightarrow$ is read as 'does not tend to'.

Next we prove that f is continuous at irrational points. Let c be an irrational number in $\mathbb{R}$. We must prove that

for each $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|f(x) - f(c)| < \varepsilon, \quad \text{for all } x \text{ with } |x - c| < \delta. \tag{14}$$

Let $\varepsilon > 0$ be given. Since c is irrational, we have $f(c) = 0$. Also, $f(x) \geq 0$ for all x in $\mathbb{R}$, so statement (14) can be rewritten as

$$f(x) < \varepsilon, \quad \text{for all } x \text{ with } |x - c| < \delta. \tag{15}$$

Note that Strategy F3 cannot be used here, since f is not a polynomial function.

To obtain a value of δ such that statement (15) holds, we first choose a positive integer N such that $1/N < \varepsilon$. Then we let S_N denote the set of rationals p/q in the interval $(c-1, c+1)$ such that $0 < q \leq N$. There are only finitely many elements of the set S_N, and $c \notin S_N$ because c is irrational. Thus the number

$$\delta = \min\{|x - c| : x \in S_N\}$$

exists and is positive. Therefore the open interval $(c - \delta, c + \delta)$ contains *no* rationals p/q with $0 < q \leq N$.

Hence if $|x - c| < \delta$, then

either $\quad x$ is irrational, so $f(x) = 0 < \varepsilon$,

or $\quad x = p/q$ with $q > N$, so $f(x) = 1/q < 1/N < \varepsilon$.

In either case $f(x) < \varepsilon$, so we have succeeded in choosing $\delta > 0$ such that statement (15) holds. Hence f is continuous at c. ■

In view of the strange properties of the Riemann function, it is natural to ask whether a function can also be found which is continuous at each rational point of $\mathbb{R}$ and discontinuous at each irrational point. It can be shown that no such function exists, but we do not prove this here.

The blancmange function

Our next function is in some ways even more unusual than the Dirichlet and the Riemann functions. To construct this function, we start with the **sawtooth function** illustrated in Figure 18:

$$s(x) = \begin{cases} x - \lfloor x \rfloor, & \text{if } 0 \leq x - \lfloor x \rfloor \leq \frac{1}{2}, \\ 1 - (x - \lfloor x \rfloor), & \text{if } \frac{1}{2} < x - \lfloor x \rfloor < 1, \end{cases}$$

where $\lfloor x \rfloor$ is the integer part function.

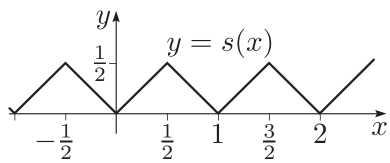


Figure 18 The sawtooth function

The **blancmange function** B is obtained by forming an infinite series of functions related to s:

$$\begin{aligned} B(x) &= s(x) + \tfrac{1}{2}s(2x) + \tfrac{1}{4}s(4x) + \tfrac{1}{8}s(8x) + \cdots \\ &= \sum_{n=0}^{\infty} \frac{1}{2^n} s(2^n x). \end{aligned}$$

Teiji Takagi

The properties of this function were first studied by the Japanese mathematician Teiji Takagi (1875–1960) in 1903. The name 'blancmange function' was used by the English mathematician David Tall (1941–) in the 1980s.

For example, to evaluate $B\left(\frac{1}{4}\right)$ we find the sum of the corresponding series:

$$\begin{aligned} B\left(\tfrac{1}{4}\right) &= s\left(\tfrac{1}{4}\right) + \tfrac{1}{2}s\left(\tfrac{1}{2}\right) + \tfrac{1}{4}s(1) + \tfrac{1}{8}s(2) + \cdots \\ &= \tfrac{1}{4} + \tfrac{1}{2} \times \tfrac{1}{2} + \tfrac{1}{4} \times 0 + \tfrac{1}{8} \times 0 + \cdots \\ &= \tfrac{1}{2}. \end{aligned}$$

In this case, the series has only finitely many non-zero terms, but for some x the series for $B(x)$ has infinitely many non-zero terms. However, since $0 \leq s(x) \leq \frac{1}{2}$, for $x \in \mathbb{R}$, we have

$$0 \leq \frac{1}{2^n}s(2^n x) \leq \frac{1}{2^{n+1}}, \quad \text{for } x \in \mathbb{R} \text{ and } n \in \mathbb{N},$$

so the series defining $B(x)$ is convergent for each $x \in \mathbb{R}$, by the Comparison Test for series; see Subsection 2.1 of Unit D3.

To picture the graph of the blancmange function, we consider the graphs of several successive partial sum functions of the series for B, with domains restricted to $[0, 1]$; see Figure 19. (In each case the graph of the previous partial sum function is in light dashes and the function being added is in heavy dashes. Thus the solid line is the sum of the graphs shown by dashed lines.)

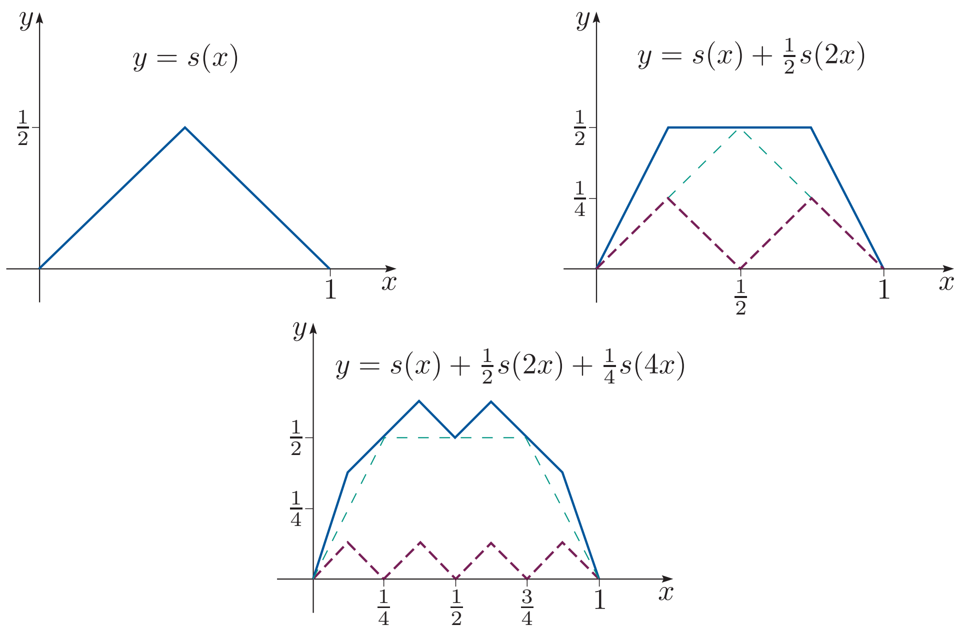


Figure 19 Successive steps in the construction of the blancmange function

The sum function B has the graph shown in Figure 20.

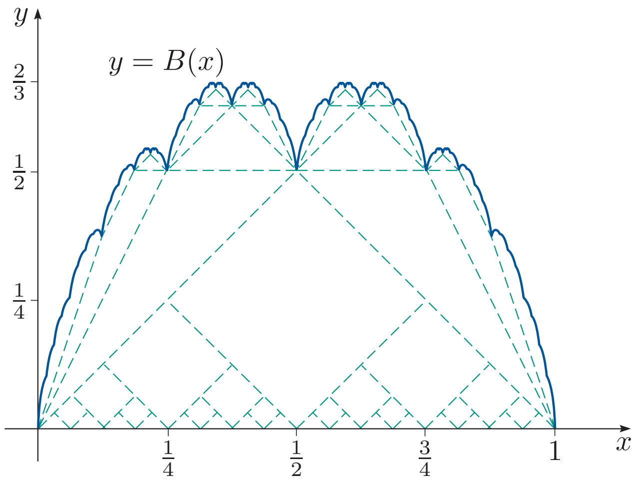


Figure 20 The blancmange function

The graph of B is very irregular, in the sense that it oscillates rapidly up and down, and does not appear to be smooth at any point. In fact, in Unit F2 you will see that the function B is nowhere differentiable! However, it does seem that the function B is continuous, and we can show that this is true.

Theorem F17

The blancmange function is continuous.

Proof We use the ε-δ definition of continuity.

Let $c \in \mathbb{R}$. We want to show that

for each $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|B(x) - B(c)| < \varepsilon, \quad \text{for all } x \text{ with } |x - c| < \delta. \tag{16}$$

Let $\varepsilon > 0$ be given. We first write

$$B(x) - B(c) = \sum_{n=0}^{\infty} \frac{1}{2^n} \left(s(2^n x) - s(2^n c)\right).$$

Hence, by the infinite form of the Triangle Inequality (see Theorem D35 in Subsection 3.1 of Unit D3),

$$|B(x) - B(c)| \leq \sum_{n=0}^{\infty} \frac{1}{2^n} |s(2^n x) - s(2^n c)|. \tag{17}$$

For all x and c, and $n = 0, 1, 2, \ldots$, both $s(2^n x)$ and $s(2^n c)$ lie in the interval $\left[0, \frac{1}{2}\right]$, so

$$|s(2^n x) - s(2^n c)| \leq \tfrac{1}{2}, \quad \text{for } n = 0, 1, 2, \ldots. \tag{18}$$

Now we choose an integer N such that $1/2^N < \frac{1}{2}\varepsilon$ and consider the 'tail' of the series in inequality (17), starting from the term $n = N$.

Such an N exists because $(1/2^n)$ is a basic null sequence. Splitting the series into two different parts will enable us to use different methods for each part. We bound each part by $\frac{1}{2}\varepsilon$ to give an overall bound of ε.

By inequality (18), we have

$$\begin{aligned}\sum_{n=N}^{\infty} \frac{1}{2^n}|s(2^n x) - s(2^n c)| &\leq \frac{1}{2}\sum_{n=N}^{\infty}\frac{1}{2^n} \\ &= \frac{1}{2}\left(\frac{1/2^N}{1-1/2}\right) = \frac{1}{2^N} < \tfrac{1}{2}\varepsilon. \qquad (19)\end{aligned}$$

Here we have used the fact that $\displaystyle\sum_{n=N}^{\infty}\frac{1}{2^n}$ is a geometric series with $a = 1/2^N$ and $r = 1/2$, so $\displaystyle\sum_{n=N}^{\infty}\frac{1}{2^n} = \frac{1/2^N}{1-1/2}$.

Next we consider the rest of this series. Each of the functions

$$x \longmapsto s(2^n x), \quad n = 0, 1, \ldots, N-1,$$

is continuous.

We omit the proof of this – you may like to check it for yourself.

Therefore, for each $n = 0, 1, \ldots, N-1$, there is a positive number δ_n such that

$$|s(2^n x) - s(2^n c)| < \tfrac{1}{4}\varepsilon, \quad \text{for all } x \text{ with } |x - c| < \delta_n.$$

You will see a bit later in the argument why it makes sense to choose $\frac{1}{4}\varepsilon$ here.

Thus if $\delta = \min\{\delta_0, \delta_1, \ldots, \delta_{N-1}\}$ and $|x - c| < \delta$, then

$$\sum_{n=0}^{N-1}\frac{1}{2^n}|s(2^n x) - s(2^n c)| < \sum_{n=0}^{N-1}\frac{1}{2^n}\left(\tfrac{1}{4}\varepsilon\right) < 2 \times \tfrac{1}{4}\varepsilon = \tfrac{1}{2}\varepsilon.$$

Here we have used the fact that $\displaystyle\sum_{n=0}^{\infty}\frac{1}{2^n} = 2$.

Combining this inequality with inequalities (19) and (17), we obtain statement (16) with this choice of δ. Hence B is continuous at the point c. ■

The blancmange function is very irregular, but it exhibits patterns known as 'self-similarity'. However closely you look at the graph, you can see 'mini-blancmanges' growing on it everywhere. The existence of these mini-blancmanges can be explained by rewriting the series defining B:

$$\begin{aligned}B(x) &= s(x) + \tfrac{1}{2}s(2x) + \tfrac{1}{4}s(4x) + \tfrac{1}{8}s(8x) + \cdots \\ &= s(x) + \tfrac{1}{2}\left(s(2x) + \tfrac{1}{2}s(4x) + \tfrac{1}{4}s(8x) + \cdots\right) \\ &= s(x) + \tfrac{1}{2}B(2x).\end{aligned}$$

The graph of the function $x \longmapsto \frac{1}{2}B(2x)$ is just the graph of B scaled by the factor $\frac{1}{2}$ in both x- and y-directions. Hence the graph of B is the graph of s with a (sheared) $\frac{1}{2}$-size blancmange growing on each sloping line segment. (You met shears in Unit C3 *Linear Transformations*.) Smaller mini-blancmanges can be explained in a similar manner. This is illustrated in Figure 21.

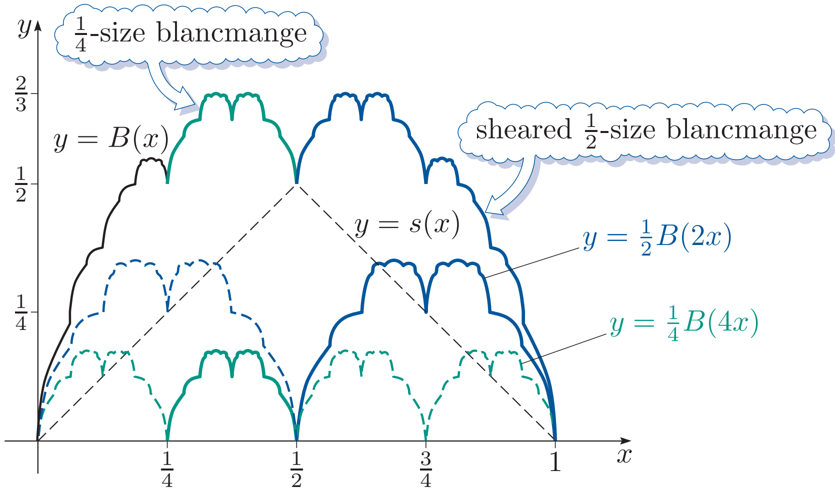


Figure 21 The self-similarity of the blancmange function

Such irregular sets, which display self-similarity, are studied in detail in the subject known as 'fractals'.

3.3 Limits and other asymptotic behaviour

In Sections 1 and 2 we defined limits and other types of asymptotic behaviour using a sequential approach. Each of these concepts can also be defined in a way that is analogous to the ε-δ definition of continuity. For example, we can define the concept of a limit as follows.

Definition

Let f be a function defined on a punctured neighbourhood $N_r(c)$ of c. Then $f(x)$ **tends to the limit l as x tends to c** if

for each $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|f(x) - l| < \varepsilon, \quad \text{for all } x \text{ with } 0 < |x - c| < \delta.$$

As before, we write

$$\lim_{x \to c} f(x) = l \quad \text{or} \quad f(x) \to l \text{ as } x \to c.$$

Remarks

1. This definition is very similar to the ε-δ definition of continuity, except that $f(c)$ is replaced by l and $|x-c| < \delta$ is replaced by $0 < |x-c| < \delta$. This reflects the fact that when we try to find the limit of $f(x)$ as x tends to c, the value of $f(x)$ at $x = c$ is not relevant and indeed may not be defined.
2. The proof that the above definition is equivalent to the sequential definition of a limit is similar to the proof of Theorem F14.
3. Analogous definitions can be constructed for one-sided limits and for all the types of asymptotic behaviour you have met.

   For example, $\lim_{x\to\infty} f(x) = l$ if

   $$\text{for each } \varepsilon > 0, \text{ there exists } K \in \mathbb{R} \text{ such that}$$
   $$|f(x) - l| < \varepsilon, \quad \text{for all } x \text{ with } x > K.$$

Karl Weierstrass

A brief history of the limit concept

Recent historical studies have shown that, contrary to what was previously believed, Isaac Newton (1642–1727) had a good grasp of the limit concept, recognising that limits provide a secure foundation for calculus. Jean le Rond d'Alembert (1717–1783) in his article in the French *Encyclopédie* (published between 1751 and 1765) provided a definition of limit very close to that of Newton: a bound that could be approached as closely as one chose. However, because d'Alembert, like Newton, worked with examples that were primarily geometric, there was no need to consider quantities that might oscillate from one side of a limit to the other.

In 1821 Augustin-Louis Cauchy (1789–1857), in his *Cours d'Analyse*, provided a definition of a limit which combined the same ideas as that of d'Alembert: the existence of a fixed value and the possibility of approaching it as closely as one wishes. Although Cauchy was the first to use ε-δ arguments in his proofs, he never gave an explicit ε-δ definition of a limit. Unknown to Cauchy, in 1817 the Bohemian theologian and mathematician Bernard Bolzano (1781–1848) had already introduced a rigorous ε-δ definition of a limit, but his work was not well known and had only indirect influence on later developments. The ε-δ definition used today was first formalised by the German mathematician Karl Weierstrass (1815–1897) in his lectures in Berlin in the 1860s.

The next worked exercise gives an example of a limit evaluated using the ε-δ definition. (In fact, for this example, the sequential definition is easier to use; see Worked Exercise F1(a).)

Worked Exercise F6

Use the ε-δ definition of a limit to evaluate

$$\lim_{x\to 2} \frac{x^2-4}{x-2}.$$

Solution

The domain of $f(x) = (x^2-4)/(x-2)$ is $\mathbb{R} - \{2\}$, so f is defined on each punctured neighbourhood of 2. Also,

$$f(x) = \frac{x^2-4}{x-2} = \frac{(x-2)(x+2)}{x-2} = x+2, \quad \text{for } x \neq 2.$$

Limits of this type with a factor that can be cancelled often arise in calculations.

This formula suggests that $\lim_{x\to 2} f(x) = 2+2 = 4$, so we must prove that

for each $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|f(x) - 4| < \varepsilon, \quad \text{for all } x \text{ with } 0 < |x-2| < \delta.$$

But $|f(x) - 4| = |x+2-4| = |x-2|$, for $x \neq 2$, so the above statement is true if we choose $\delta = \varepsilon$. Hence

$$\lim_{x\to 2} \frac{x^2-4}{x-2} = 4.$$

This is similar to a proof, using Strategy F3, that the function $x \longmapsto x+2$ is continuous at the point 2.

Exercise F12

Use the ε-δ definition of a limit to evaluate

$$\lim_{x\to 1} \frac{2x^3+3x-5}{x-1}.$$

Hint: Use the fact that $2x^3+3x-5 = (x-1)(2x^2+2x+5)$ and follow Strategy F3 for using the ε-δ definition of continuity.

4 Uniform continuity

In Section 3 you met an alternative approach to continuity, based on the ε-δ definition. In this section you will see how this approach can be used to describe a stronger notion of continuity, which plays a key role in our later work on the integration of continuous functions in Unit F3 *Integration*.

4.1 What is uniform continuity?

The ε-δ definition of continuity states that a function f is continuous at a point c in an interval I in the domain of f if

$$\text{for each } \varepsilon > 0, \text{ there exists } \delta > 0 \text{ such that}$$
$$|f(x) - f(c)| < \varepsilon, \quad \text{for all } x \in I \text{ with } |x - c| < \delta.$$

In this definition we cannot expect that, for a given positive number ε, the same positive number δ will serve equally well for each point c in I. Sometimes, however, this does happen, in which case:

$$\text{for each } \varepsilon > 0, \text{ there exists } \delta > 0 \text{ such that}$$
$$|f(x) - f(c)| < \varepsilon, \quad \text{for all } x, c \in I \text{ with } |x - c| < \delta.$$

Informally, for all c in the interval I, as x gets closer to c, the values $f(x)$ get closer to $f(c)$ at least as quickly as some uniform rate. We make the following definition.

Definition

A function f defined on an interval I is **uniformly continuous** on I if

$$\text{for each } \varepsilon > 0, \text{ there exists } \delta > 0 \text{ such that}$$
$$|f(x) - f(y)| < \varepsilon, \quad \text{for all } x, y \in I \text{ with } |x - y| < \delta. \tag{20}$$

Remarks

1. In this definition we have used the variables x and y, rather than x and c, to indicate that these two variables are of equal standing. Notice that uniform continuity is defined on an *interval*: it is meaningless to say that a function is uniformly continuous at a *point*.
2. We say that c is an **interior point** of an interval I if c is not an endpoint of I. It follows from the above definition that if f is uniformly continuous on an interval I, then f is continuous at each interior point of I. At an endpoint of I, the function can be discontinuous because of its behaviour *outside* I.
3. If f is uniformly continuous on an interval I, then f is uniformly continuous on any subinterval of I.

Worked Exercise F7

Prove from the definition that $f(x) = x^2$ is uniformly continuous on $I = [-4, 4]$.

Solution

Let $\varepsilon > 0$ be given. We have

$$f(x) - f(y) = x^2 - y^2 = (x + y)(x - y).$$

Hence, for $x, y \in [-4, 4]$,

$$\begin{aligned} |f(x) - f(y)| &= |x + y|\,|x - y| \\ &\leq (|x| + |y|)|x - y| \quad \text{(by the Triangle Inequality)} \\ &\leq 8|x - y|, \end{aligned}$$

since $|x| \leq 4$ and $|y| \leq 4$.

Thus, if we choose $\delta = \frac{1}{8}\varepsilon$, then whenever $x, y \in [-4, 4]$ and $|x - y| < \delta$, we have

$$\begin{aligned} |f(x) - f(y)| &\leq 8|x - y| \\ &< 8 \times \tfrac{1}{8}\varepsilon = \varepsilon. \end{aligned}$$

Hence f is uniformly continuous on $[-4, 4]$.

In Exercise F13(b) you will see that $f(x) = x^2$ is *not* uniformly continuous on $\mathbb{R}$. It is, however, uniformly continuous on all bounded closed intervals, as you will see in Subsection 4.2.

In the next worked exercise we consider the function $f(x) = 1/x$. In Unit D4 you saw that this function is continuous on its domain, and hence on $(0, 1]$. We now show that it is not *uniformly* continuous on $(0, 1]$.

Worked Exercise F8

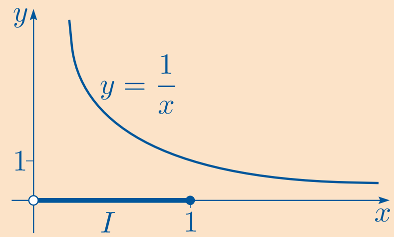
Prove that $f(x) = 1/x$ is not uniformly continuous on $I = (0, 1]$.

Solution

To prove that a function is *not* uniformly continuous, we need to show that the negation of statement (20) holds.

We have to find $\varepsilon > 0$ such that, no matter which $\delta > 0$ is chosen, there are points x and y in I with $|x - y| < \delta$ and $|f(x) - f(y)| \geq \varepsilon$.

We sketch the graph of the function f.



The graph suggests that, for any positive δ, we should take x and y near 0 because then x and y are close together but $f(x)$ and $f(y)$ can be far apart.

We try $x = \frac{1}{2}\delta$ and $y = \delta$, where $0 < \delta < 1$. Then

$$|x - y| = \left|\tfrac{1}{2}\delta - \delta\right| = \tfrac{1}{2}\delta < \delta$$

and

$$|f(x) - f(y)| = \left|\frac{1}{\frac{1}{2}\delta} - \frac{1}{\delta}\right| = \frac{2}{\delta} - \frac{1}{\delta} = \frac{1}{\delta} > 1.$$

Hence the negation of statement (20) holds with $\varepsilon = 1$, so f is not uniformly continuous on I.

The reasoning in this solution is quite subtle and can be tricky to apply, so it is useful to reformulate what it means to say that a function is *not* uniformly continuous on an interval. Roughly speaking, this happens if you can find pairs of points in the interval, as close together as you like, whose images are *not* close together. We prove the following result, illustrated in Figure 22.

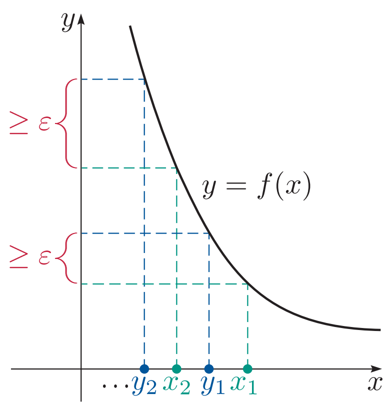


Figure 22 A function which is not uniformly continuous

Theorem F18

Let the function f be defined on an interval I. Then f is not uniformly continuous on I if and only if there exist two sequences (x_n) and (y_n) in I, and $\varepsilon > 0$, such that

1. $|x_n - y_n| \to 0$ as $n \to \infty$
2. $|f(x_n) - f(y_n)| \geq \varepsilon$, for $n = 1, 2, \dots$.

Proof We start by proving the *only if* statement.

First suppose that f is not uniformly continuous on I. Then there exists $\varepsilon > 0$ such that for all $\delta > 0$ there are points x and y in I with

$$|x - y| < \delta \quad \text{and} \quad |f(x) - f(y)| \geq \varepsilon.$$

Applying this fact with $\delta = 1$, $\delta = \frac{1}{2}$, $\delta = \frac{1}{3}$, and so on, we obtain sequences (x_n) and (y_n) in I such that

$$|x_n - y_n| < \frac{1}{n} \quad \text{and} \quad |f(x_n) - f(y_n)| \geq \varepsilon, \quad \text{for } n = 1, 2, \ldots.$$

Thus statements 1 and 2 both hold.

On the other hand, suppose that there exist sequences (x_n) and (y_n) in I, and $\varepsilon > 0$, such that statements 1 and 2 hold.

We use a proof by contradiction to show that f is not uniformly continuous on I.

If f *is* uniformly continuous on I, then there exists $\delta > 0$ such that

$$|f(x) - f(y)| < \varepsilon, \quad \text{for all } x, y \in I \text{ with } |x - y| < \delta.$$

But $|x_n - y_n| < \delta$, for $n > N$ say, by statement 1, so

$$|f(x_n) - f(y_n)| < \varepsilon, \quad \text{for } n > N,$$

contradicting statement 2. Thus f is not uniformly continuous on I. ■

Theorem F18 gives us the second part of the following strategy; the first part is an elaboration of the definition of uniform continuity.

Strategy F4

- To prove that a function f is uniformly continuous on an interval I, find an expression for $\delta > 0$ in terms of a given $\varepsilon > 0$ such that

  $$|f(x) - f(y)| < \varepsilon, \quad \text{for all } x, y \in I \text{ with } |x - y| < \delta.$$

- To prove that a function f is *not* uniformly continuous on an interval I, find two sequences (x_n) and (y_n) in I, and $\varepsilon > 0$, such that

  $$|x_n - y_n| \to 0 \text{ as } n \to \infty, \text{ and}$$
  $$|f(x_n) - f(y_n)| \geq \varepsilon, \quad \text{for } n = 1, 2, \ldots.$$

When using part 2 of Strategy F4, you should aim to choose the terms x_n and y_n close together at points of I where the graph of f is steep; see Exercise F13(b), for example.

We could have applied part 2 of Strategy F4 in Worked Exercise F8 by taking $x_n = 1/(2n)$ and $y_n = 1/n$, for $n = 1, 2, \ldots$, since (x_n) and (y_n) lie in I:

$$|x_n - y_n| = \left| \frac{1}{2n} - \frac{1}{n} \right| = \frac{1}{2n} \to 0 \text{ as } n \to \infty$$

and, for $n = 1, 2, \ldots$,

$$|f(x_n) - f(y_n)| = \left| \frac{1}{x_n} - \frac{1}{y_n} \right| = \left| \frac{1}{1/(2n)} - \frac{1}{1/n} \right| = n \geq 1.$$

Thus, by taking $\varepsilon = 1$ in part 2 of Strategy F4, we deduce that f is not uniformly continuous on I.

Try the next exercise using Strategy F4.

Exercise F13

(a) Prove that $f(x) = x^3$ is uniformly continuous on $I = [-2, 2]$.

(b) Prove that $f(x) = x^2$ is not uniformly continuous on $I = \mathbb{R}$.

Hint: In part (b), take $x_n = n + 1/n$ and $y_n = n$, for $n = 1, 2, \ldots$.

4.2 A condition that ensures uniform continuity

Checking uniform continuity from the definition can be complicated. However, we can often deduce uniform continuity in a straightforward way from the following fundamental result. Like the Intermediate Value Theorem and the Extreme Value Theorem that you met in Section 3 of Unit D4, this result is another illustration of the fact that continuous functions on bounded closed intervals have particularly nice properties.

Theorem F19

If the function f is continuous on a bounded closed interval $[a, b]$, then f is uniformly continuous on $[a, b]$.

You saw in Worked Exercise F7 that the function $f(x) = x^2$ is uniformly continuous on $[-4, 4]$. In fact this can be deduced immediately from Theorem F19, as follows. The function $f(x) = x^2$ is continuous on the whole of $\mathbb{R}$, and hence on any bounded closed interval, so it must be uniformly continuous on *any* bounded closed interval, by Theorem F19. However, $f(x) = x^2$ is not *uniformly* continuous on the set $\mathbb{R}$, as you saw in Exercise F13(b): the image values of points near to a point c approach $f(c)$ more and more slowly as the point c gets larger and larger.

Exercise F14

Use Theorem F19 to prove that $f(x) = x^3$ is uniformly continuous on $I = [-2, 2]$.

Proof of Theorem F19 We assume that f is continuous on $[a, b]$ but *not* uniformly continuous on $[a, b]$, and deduce a contradiction using the bisection method.

You met the bisection method in the proofs of the Intermediate Value Theorem and the Extreme Value Theorem in Section 3 of Unit D4.

Since we are assuming that f is not uniformly continuous on $[a, b]$, it follows from Theorem F18 that there exist sequences (x_n) and (y_n) in $[a, b]$ and $\varepsilon > 0$, such that

$$|x_n - y_n| \to 0 \text{ as } n \to \infty, \text{ and} \tag{21}$$

$$|f(x_n) - f(y_n)| \geq \varepsilon, \quad \text{for } n = 1, 2, \ldots. \tag{22}$$

Let $a_0 = a$, $b_0 = b$ and $p = \frac{1}{2}(a_0 + b_0)$. Then at least one of $[a_0, p]$ or $[p, b_0]$ must contain terms x_n for infinitely many $n \in \mathbb{N}$. We denote this interval by $[a_1, b_1]$, choosing either interval if both contain infinitely many terms x_n. Thus we have:

1. $[a_1, b_1] \subseteq [a_0, b_0]$
2. $b_1 - a_1 = \frac{1}{2}(b_0 - a_0)$
3. $[a_1, b_1]$ contains terms x_n for infinitely many $n \in \mathbb{N}$.

Now we repeat this process, bisecting $[a_1, b_1]$ to obtain $[a_2, b_2]$, and so on. This gives a sequence of closed intervals

$$[a_k, b_k], \quad k = 0, 1, 2, \ldots,$$

such that the following properties hold for $k = 0, 1, 2, \ldots$

1. $[a_{k+1}, b_{k+1}] \subseteq [a_k, b_k]$
2. $b_k - a_k = \left(\frac{1}{2}\right)^k (b_0 - a_0)$
3. $[a_k, b_k]$ contains terms x_n for infinitely many $n \in \mathbb{N}$.

We use k here to avoid n having two meanings.

Now property 1 implies that (a_k) is increasing and bounded above by b_0. Hence (a_k) is convergent by the Monotone Convergence Theorem (see Section 5 of Unit D2).

Moreover, if we let $\lim\limits_{k\to\infty} a_k = c$, then it follows that $\lim\limits_{k\to\infty} b_k = c$ also, by property 2 and the Combination Rules for sequences, and by property 3, the sequence (x_n) contains a subsequence (x_{n_k}) such that

$$x_{n_k} \in [a_k, b_k], \quad \text{for } k = 0, 1, 2, \ldots.$$

Thus $\lim\limits_{k\to\infty} x_{n_k} = c$, by the Squeeze Rule for sequences, so

$$y_{n_k} = (y_{n_k} - x_{n_k}) + x_{n_k} \to 0 + c = c,$$

by statement (21). It now follows from the continuity of f at the point c that

$$\lim_{k\to\infty} f(x_{n_k}) = f(c) \quad \text{and} \quad \lim_{k\to\infty} f(y_{n_k}) = f(c),$$

but this contradicts statement (22). Hence our original assumption that f is continuous but not uniformly continuous on $[a, b]$ must be false. This completes the proof. ■

As part of this proof, we showed that any sequence (x_n) in $[a, b]$ contains a convergent subsequence (x_{n_k}). This remarkable result, which is of importance in many parts of analysis, is called the *Bolzano–Weierstrass Theorem*. It can be stated as follows.

Theorem F20 Bolzano–Weierstrass Theorem

Any bounded sequence has a convergent subsequence.

Summary

In this unit you have studied the limiting behaviour of a real function f near a point c, using the concept of a punctured neighbourhood of a point c. You have seen that $f(x)$ has limit $f(c)$ as x tends to c precisely when f is continuous at the point c. You have also studied the limiting behaviour of f as x tends to infinity, and seen what it means for the function values $f(x)$ to tend to infinity as x tends to c. You have learnt how to use a variety of rules to help you analyse such asymptotic behaviour and to evaluate different limits.

You have also met a new definition of continuity: the ε-δ definition of continuity. You have seen that this is equivalent to the sequential definition of continuity that you studied in Unit D4 and, although it is more abstract than the sequential definition, it can be easier to use in certain situations. In particular, you have used both definitions to prove different properties of some interesting functions: the Dirichlet function, the Riemann function and the blancmange function.

Finally you have learnt what it means for a function to be *uniformly continuous* on an interval. You have seen how to give a direct proof that a function has this property and met the result that a function that is continuous on a bounded closed interval is uniformly continuous.

As you continue your studies, you will see that the ideas you have met in this unit play a key role in the foundations of calculus. In particular, limits are fundamental in the theories of differentiation and integration, and uniform continuity is used to show that a function that is continuous on a bounded closed interval can be integrated.

Learning outcomes

After working through this unit, you should be able to:

- understand the statement $\lim_{x \to c} f(x) = l$, or $f(x) \to l$ as $x \to c$
- appreciate the relationship between limits of a function and continuity
- use the Combination Rules, Composition Rule and Squeeze Rule to calculate limits of functions
- understand one-sided limits, and the statements $\lim_{x \to c^+} f(x) = l$ and $\lim_{x \to c^-} f(x) = l$
- use certain basic limits to find other limits
- understand the *asymptotic behaviour* of functions, and the statements $f(x) \to \infty$ as $x \to c$ (or c^+ or c^-) and $f(x) \to l$ as $x \to \infty$ (or $-\infty$)
- use the Reciprocal Rule, the Combination Rules, the Squeeze Rule and the Composition Rule to determine the asymptotic behaviour of functions
- understand and use the ε–δ definition of continuity
- describe functions which are highly discontinuous, or continuous but very irregular
- understand and use the ε–δ definition of a limit
- understand the definition of *uniform continuity* and use it in simple cases
- know and use various conditions for uniform continuity.

Solutions to exercises

Solution to Exercise F1

(a) The function

$$f(x) = \frac{x^2 + x}{x}$$

has domain $\mathbb{R} - \{0\}$, so f is defined on any punctured neighbourhood of 0. Also,

$$f(x) = \frac{x(x+1)}{x} = x + 1, \quad \text{for } x \neq 0.$$

Thus if (x_n) lies in $\mathbb{R} - \{0\}$ and $x_n \to 0$, then

$$f(x_n) = x_n + 1 \to 1.$$

Hence

$$\lim_{x \to 0} \frac{x^2 + x}{x} = 1.$$

(b) The domain of $f(x) = \lfloor x \rfloor$ is $\mathbb{R}$, so f is defined on any punctured neighbourhood of 1. Now,

$$f(x) = 0, \quad \text{for } 0 \leq x < 1,$$

and

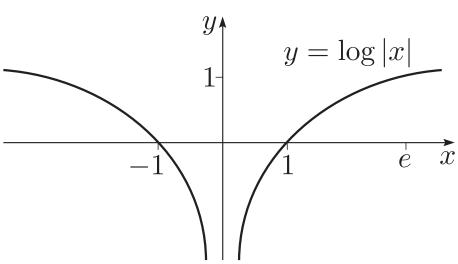
$$f(x) = 1, \quad \text{for } 1 \leq x < 2.$$

The two sequences $(1 + 1/n)$ and $(1 - 1/n)$ both tend to 1, and have terms lying in $\mathbb{R} - \{1\}$. Also,

$$\lim_{n \to \infty} f(1 + 1/n) = 1 \quad \text{but} \quad \lim_{n \to \infty} f(1 - 1/n) = 0.$$

Hence $\lim_{x \to 1} \lfloor x \rfloor$ does not exist, by the first part of Strategy F1.

(c) The function $f(x) = \log |x|$ has domain $\mathbb{R} - \{0\}$, so it is defined on any punctured neighbourhood of 0, and its graph is as follows. (This is included for interest – you are not expected to have sketched the graph as part of your solution.)



We consider the null sequence $x_n = 1/n$. Then

$$f(x_n) = \log |x_n| = \log(1/n) = -\log n.$$

Now $\log n \to \infty$ as $n \to \infty$, so $f(x_n) \to -\infty$ as $n \to \infty$.

Hence $\lim_{x \to 0} \log |x|$ does not exist, by the second part of Strategy F1.

Solution to Exercise F2

(a) The domain of the continuous function $f(x) = \sqrt{x}$ is the interval $[0, \infty)$, so f is defined on the open interval $(0, \infty)$, which contains 2. Hence, by Theorem F2, we have

$$\lim_{x \to 2} \sqrt{x} = \sqrt{2}.$$

(b) Since $\sin x > 0$ for $0 < x < \pi$, the function $f(x) = \sqrt{\sin x}$ is defined on the open interval $(0, \pi)$, which contains $\pi/2$, and is continuous, by the Composition Rule for continuous functions. Hence, by Theorem F2, we have

$$\lim_{x \to \pi/2} \sqrt{\sin x} = \sqrt{\sin(\pi/2)} = 1.$$

(c) The function $f(x) = e^x/(1 + x)$ is defined on the open interval $(-1, \infty)$, which contains 1, and is continuous, by the Quotient Rule for continuous functions. Hence, by Theorem F2, we have

$$\lim_{x \to 1} \frac{e^x}{1 + x} = \frac{e^1}{1 + 1} = \tfrac{1}{2}e.$$

Solution to Exercise F3

(a) First we write

$$\frac{\sin x}{2x + x^2} = \frac{\sin x}{x(2 + x)} = \left(\frac{\sin x}{x}\right)\left(\frac{1}{2 + x}\right).$$

Since

$$\lim_{x \to 0} \frac{\sin x}{x} = 1, \quad \text{(by Theorem F1)}$$

and

$$\lim_{x \to 0} \frac{1}{2 + x} = \tfrac{1}{2},$$

because $\dfrac{1}{2 + x}$ is continuous at 0, we deduce by the

Product Rule for limits that

$$\lim_{x\to 0}\frac{\sin x}{2x+x^2} = 1\times\tfrac{1}{2} = \tfrac{1}{2}.$$

(b) We can write

$$\frac{\sin(\sin x)}{\sin x} = g(f(x)),$$

where $f(x) = \sin x$ and $g(x) = \dfrac{\sin x}{x}$.

Substituting $u = f(x) = \sin x$ and using the fact that u is continuous at 0, we have

$$u = \sin x \to \sin 0 = 0 \quad \text{as } x \to 0$$

and

$$g(u) = \frac{\sin u}{u} \to 1 \quad \text{as } u \to 0.$$

The first proviso to the Composition Rule holds, because $f(x) = \sin x \neq 0$ in $N_1(0)$, for example.

Thus, by the Composition Rule,

$$g(f(x)) = \frac{\sin(\sin x)}{\sin x} \to 1 \quad \text{as } x \to 0.$$

(c) We can write

$$\left(\frac{x}{\sin x}\right)^{1/2} = g(f(x)),$$

where $f(x) = \dfrac{\sin x}{x}$ and $g(x) = 1/x^{1/2}$.

Substituting $u = f(x) = \dfrac{\sin x}{x}$, we have

$$u = \frac{\sin x}{x} \to 1 \quad \text{as } x \to 0,$$
$$g(u) = 1/u^{1/2} \to 1/1^{1/2} = 1 \quad \text{as } u \to 1,$$

since g is continuous at 1, which also tells us that the second proviso to the Composition Rule holds.

Thus, by the Composition Rule,

$$g(f(x)) = \left(\frac{x}{\sin x}\right)^{1/2} \to 1 \quad \text{as } x \to 0.$$

(d) Using the hint, we obtain

$$\begin{aligned}\frac{1-\cos x}{x} &= \frac{2\sin^2\left(\frac{1}{2}x\right)}{x}\\ &= \frac{\sin\left(\frac{1}{2}x\right)}{\frac{1}{2}x}\times\sin\left(\tfrac{1}{2}x\right).\end{aligned}$$

Now,

$$\lim_{x\to 0}\frac{\sin\left(\frac{1}{2}x\right)}{\frac{1}{2}x} = 1,$$

by Worked Exercise F3(a), and

$$\lim_{x\to 0}\sin\left(\tfrac{1}{2}x\right) = 0,$$

by the continuity of the function $x \longmapsto \sin\left(\frac{1}{2}x\right)$.

Thus, by the Product Rule,

$$\lim_{x\to 0}\frac{1-\cos x}{x} = 1\times 0 = 0.$$

Solution to Exercise F4

(a) Since

$$1 + x \le e^x \le \frac{1}{1-x}, \quad \text{for } |x| < 1,$$

we have

$$x \le e^x - 1 \le \frac{1}{1-x} - 1 = \frac{x}{1-x}, \quad \text{for } |x| < 1.$$

Thus if $0 < x < 1$, then

$$1 \le \frac{e^x-1}{x} \le \frac{1}{1-x} = 1 + \frac{x}{1-x} = 1 + \frac{|x|}{1-x},$$

and if $-1 < x < 0$, then

$$1 \ge \frac{e^x-1}{x} \ge \frac{1}{1-x} = 1 + \frac{x}{1-x} = 1 - \frac{|x|}{1-x}.$$

Since we also have

$$1 - \frac{|x|}{1-x} < 1 \quad \text{when } 0 < x < 1$$

and

$$1 + \frac{|x|}{1-x} > 1 \quad \text{when } -1 < x < 0,$$

we deduce that the inequalities

$$1 - \frac{|x|}{1-x} \le \frac{e^x-1}{x} \le 1 + \frac{|x|}{1-x}$$

hold whenever $0 < |x| < 1$, as required.

(b) We have

$$\lim_{x\to 0}\left(1 - \frac{|x|}{1-x}\right) = 1$$

and

$$\lim_{x\to 0}\left(1 + \frac{|x|}{1-x}\right) = 1,$$

because both functions are continuous on the interval $(-1, 1)$.

Thus, by part (a) and the Squeeze Rule,

$$\lim_{x \to 0} \frac{e^x - 1}{x} = 1.$$

Solution to Exercise F5

(a) Since $\lim_{x \to 0} \frac{\sin x}{x} = 1$ and the function $x \longmapsto \sqrt{x}$ is continuous at 0, we have

$$\lim_{x \to 0^+} \frac{\sin x}{x} = 1 \quad \text{(by Theorem F7)}$$

and

$$\lim_{x \to 0^+} \sqrt{x} = 0 \quad \text{(by Theorem F8).}$$

Hence, by the Sum Rule,

$$\lim_{x \to 0^+} \left(\frac{\sin x}{x} + \sqrt{x} \right) = 1 + 0 = 1.$$

(b) We can write

$$\frac{\sin \sqrt{x}}{\sqrt{x}} = g(f(x)),$$

where $f(x) = \sqrt{x}$ and $g(x) = \dfrac{\sin x}{x}$.

Substituting $u = f(x) = \sqrt{x}$, we have

$$u = \sqrt{x} \to 0 \quad \text{as } x \to 0^+,$$

$$g(u) = \frac{\sin u}{u} \to 1 \quad \text{as } u \to 0,$$

so that

$$\lim_{u \to 0^+} \frac{\sin u}{u} = 1 \quad \text{(by Theorem F7).}$$

Moreover, $f(x) = \sqrt{x} \neq 0$ on the open interval $(0, 1)$, so the first proviso of the one-sided limit version of the Composition Rule holds.

Thus, by the Composition Rule,

$$g(f(x)) = \frac{\sin \sqrt{x}}{\sqrt{x}} \to 1 \quad \text{as } x \to 0^+.$$

Solution to Exercise F6

(a) Let $f(x) = |x|$; then $f(x) > 0$ for $x \in \mathbb{R} - \{0\}$, and

$$\lim_{x \to 0} |x| = 0,$$

since f is continuous at 0. Hence, by the Reciprocal Rule,

$$\frac{1}{f(x)} = \frac{1}{|x|} \to \infty \quad \text{as } x \to 0.$$

(b) Let $f(x) = x^3/\sin x$; then

$$f(x) = \frac{x^2}{(\sin x)/x} > 0, \quad \text{for } x \in N_\pi(0),$$

and

$$f(x) = \frac{x^2}{(\sin x)/x} \to \frac{0}{1} = 0 \quad \text{as } x \to 0,$$

by Theorem F6(a) and the Quotient Rule, since the function $x \longmapsto x^2$ is continuous at 0.

Hence, by the Reciprocal Rule,

$$\frac{1}{f(x)} = \frac{\sin x}{x^3} \to \infty \quad \text{as } x \to 0.$$

(c) Let $f(x) = x^3 - 1$; then $f(x) > 0$ for $x \in (1, \infty)$, and

$$\lim_{x \to 1^+} x^3 - 1 = 0,$$

since f is continuous at 1. Hence, by the Reciprocal Rule,

$$\frac{1}{f(x)} = \frac{1}{x^3 - 1} \to \infty \quad \text{as } x \to 1^+.$$

Solution to Exercise F7

(a) Since

$$f(x) = \frac{2x^3 + x}{x^3} = 2 + \frac{1}{x^2}, \quad \text{for } x \neq 0,$$

we deduce, by Theorem F11(b) and the Sum Rule, that

$$f(x) = 2 + \frac{1}{x^2} \to 2 + 0 = 2 \quad \text{as } x \to \infty.$$

(b) Let

$$f(x) = \frac{x^2}{2x^3 + 1};$$

then dividing both the numerator and the denominator by the dominant term x^3, we obtain

$$f(x) = \frac{1/x}{2 + 1/x^3} \to \frac{0}{2 + 0} = 0 \quad \text{as } x \to \infty,$$

by Theorem F11(b) and the Combination Rules.

Hence, by the Reciprocal Rule,

$$\frac{1}{f(x)} = \frac{2x^3+1}{x^2} \to \infty \quad \text{as } x \to \infty.$$

Solution to Exercise F8

Since $-1 \le \sin(1/x) \le 1$ for $x \ne 0$, we have

$$-\frac{1}{x} \le \frac{\sin(1/x)}{x} \le \frac{1}{x}, \quad \text{for } x \in (0,\infty).$$

Also, by Theorem F11(b) and the Multiple Rule,

$$g(x) = -\frac{1}{x} \to 0 \quad \text{as } x \to \infty$$

and

$$h(x) = \frac{1}{x} \to 0 \quad \text{as } x \to \infty.$$

Thus, by the Squeeze Rule, part (a),

$$f(x) = \frac{\sin(1/x)}{x} \to 0 \quad \text{as } x \to \infty.$$

Solution to Exercise F9

(a) By Theorem F13(b), we have

$$\frac{e^x}{x^2} \to \infty \quad \text{as } x \to \infty.$$

By Theorem F13(c), we have

$$\frac{\log x}{x^2} \to 0 \quad \text{as } x \to \infty.$$

Also, $(\log x)/x^2 > 0$ for $x > 1$, so

$$\frac{x^2}{\log x} \to \infty \quad \text{as } x \to \infty,$$

by the Reciprocal Rule.

Hence, by the Combination Rules,

$$\frac{e^x}{x^2} + \frac{3x^2}{\log x} \to \infty \quad \text{as } x \to \infty.$$

(b) Following the hint, we write

$$\frac{\log x}{e^x} = \left(\frac{\log x}{x}\right)\left(\frac{x}{e^x}\right).$$

By Theorem F13(b) and (c),

$$\frac{x}{e^x} \to 0 \quad \text{as } x \to \infty \quad \text{and} \quad \frac{\log x}{x} \to 0 \quad \text{as } x \to \infty.$$

Thus, by the Product Rule,

$$\frac{\log x}{e^x} \to 0 \times 0 = 0 \quad \text{as } x \to \infty.$$

(c) The dominant term is e^x, so we write

$$\frac{2e^x - x^2}{e^x + \log x} = \frac{2 - x^2/e^x}{1 + (\log x)/e^x}.$$

Thus, by part (b) above, Theorem F13(b) and the Combination Rules,

$$\frac{2e^x - x^2}{e^x + \log x} \to \frac{2-0}{1+0} = 2 \quad \text{as } x \to \infty.$$

Solution to Exercise F10

(a) We can write

$$\frac{e^{x^2}}{x^2} = g(f(x)),$$

where $f(x) = x^2$ and $g(x) = e^x/x$.

Substituting $u = f(x) = x^2$, we have (by Theorem F13(a) and (b))

$$u = x^2 \to \infty \quad \text{as } x \to \infty,$$
$$g(u) = \frac{e^u}{u} \to \infty \quad \text{as } u \to \infty.$$

Thus, by the Composition Rule,

$$g(f(x)) = \frac{e^{x^2}}{x^2} \to \infty \quad \text{as } x \to \infty.$$

(b) We can write

$$\log(\log x) = g(f(x)),$$

where $f(x) = \log x$ and $g(x) = \log x$.

Substituting $u = f(x) = \log x$, we have (by Theorem F13(c))

$$u = \log x \to \infty \quad \text{as } x \to \infty,$$
$$g(u) = \log u \to \infty \quad \text{as } u \to \infty.$$

Thus, by the Composition Rule,

$$g(f(x)) = \log(\log x) \to \infty \quad \text{as } x \to \infty.$$

(c) We can write

$$x\sin(1/x) = \frac{\sin(1/x)}{1/x} = g(f(x)),$$

where $f(x) = \dfrac{1}{x}$ and $g(x) = \dfrac{\sin x}{x}$.

Substituting $u = f(x) = 1/x$, we have

$$u = 1/x \to 0 \quad \text{as } x \to \infty \quad \text{(by Theorem F11(b))},$$
$$g(u) = \frac{\sin u}{u} \to 1 \quad \text{as } u \to 0 \quad \text{(by Theorem F1)}.$$

Thus, by the Composition Rule,

$$g(f(x)) = x\sin(1/x) \to 1 \text{ as } x \to \infty.$$

Solution to Exercise F11

The domain of $f(x) = x^3$ is $\mathbb{R}$.

Let $\varepsilon > 0$ be given. We want to choose $\delta > 0$, in terms of ε, such that

$$|f(x)-f(1)| < \varepsilon, \quad \text{for all } x \text{ with } |x-1| < \delta. \quad (*)$$

We follow the steps in Strategy F3.

1. First we write

$$f(x) - f(1) = x^3 - 1 = (x-1)(x^2+x+1).$$

2. Next we obtain an upper bound for $|x^2+x+1|$ when x is near 1. If $|x-1| \le 1$, then x lies in the interval $[0,2]$, so (by the Triangle Inequality)

$$\begin{aligned}|x^2+x+1| &\le |x|^2 + |x| + 1\\ &\le 2^2+2+1 = 7.\end{aligned}$$

3. Hence

$$|f(x) - f(1)| \le 7|x-1|, \quad \text{for } |x-1| \le 1.$$

So if $|x-1| < \delta$, where $0 < \delta \le 1$, then

$$|f(x)-f(1)| < 7\delta.$$

Thus, if we choose $\delta = \min\{1, \frac{1}{7}\varepsilon\}$, then

$$|f(x)-f(1)| < 7\delta \le \varepsilon,$$
for all x with $|x-1| < \delta$,

which proves statement $(*)$.

Thus f is continuous at the point 1.

Solution to Exercise F12

The domain of

$$f(x) = \frac{2x^3+3x-5}{x-1}$$

is $\mathbb{R} - \{1\}$, so f is defined on each punctured neighbourhood of 1. Also, for $x \neq 1$,

$$\begin{aligned}f(x) = \frac{2x^3+3x-5}{x-1} &= \frac{(x-1)(2x^2+2x+5)}{(x-1)}\\ &= 2x^2+2x+5.\end{aligned}$$

This suggests that

$$\lim_{x\to 1} f(x) = 2\times 1^2 + 2\times 1 + 5 = 9,$$

so we must prove that

for each $\varepsilon > 0$, there exists $\delta > 0$ such that $|f(x)-9| < \varepsilon$, for all x with $0 < |x-1| < \delta$. $(*)$

1. First we write, for $x \neq 1$,

$$\begin{aligned}f(x) - 9 &= 2x^2+2x+5-9\\ &= 2(x^2+x-2)\\ &= 2(x-1)(x+2).\end{aligned}$$

2. Next, if $|x-1| \le 1$, then x lies in the interval $[0,2]$, so (by the Triangle Inequality)

$$|2(x+2)| \le 2(|x|+2) \le 2(2+2) = 8.$$

3. Hence

$$|f(x)-9| \le 8|x-1|, \quad \text{for } 0 < |x-1| \le 1.$$

So if $0 < |x-1| < \delta$, where $0 < \delta \le 1$, then

$$|f(x)-9| < 8\delta.$$

Thus, if we choose $\delta = \min\{1, \frac{1}{8}\varepsilon\}$, then

$$|f(x)-9| < 8\delta \le \varepsilon,$$

for all x with $0 < |x-1| < \delta$,

which proves statement $(*)$.

Hence

$$\lim_{x\to 1} f(x) = 9.$$

Solution to Exercise F13

(a) We use part 1 of Strategy F4. Let $\varepsilon > 0$ be given. We have

$$\begin{aligned}f(x) - f(y) &= x^3 - y^3\\ &= (x-y)(x^2+xy+y^2),\end{aligned}$$

so for $x, y \in [-2,2]$ (by the Triangle Inequality),

$$\begin{aligned}|f(x)-f(y)| &= |x-y|\,|x^2+xy+y^2|\\ &\le (|x|^2+|x|\,|y|+|y|^2)|x-y|\\ &\le 12|x-y|,\end{aligned}$$

since $|x| \le 2$ and $|y| \le 2$.

Thus, if we choose $\delta = \frac{1}{12}\varepsilon$, then whenever $x, y \in [-2,2]$ and $|x-y| < \delta$, we have

$$|f(x)-f(y)| \le 12|x-y| < 12 \times \tfrac{1}{12}\varepsilon = \varepsilon.$$

Hence f is uniformly continuous on $[-2,2]$.

(b) Following part 2 of Strategy F4 and the hint, we take $x_n = n + 1/n$ and $y_n = n$, for $n = 1, 2, \ldots$. Both sequences lie in $I = \mathbb{R}$ and

$$\begin{aligned}|x_n - y_n| &= |(n + 1/n) - n| \\ &= 1/n \to 0 \quad \text{as } n \to \infty,\end{aligned}$$

$$\begin{aligned}|f(x_n) - f(y_n)| &= |x_n^2 - y_n^2| \\ &= (n + 1/n)^2 - n^2 \\ &= n^2 + 2n(1/n) + (1/n)^2 - n^2 \\ &= 2 + 1/n^2 \geq 2, \text{ for } n = 1, 2, \ldots.\end{aligned}$$

Thus, by taking $\varepsilon = 2$ in part 2 of Strategy F4, we deduce that f is not uniformly continuous on $\mathbb{R}$.

Solution to Exercise F14

Since $f(x) = x^3$ is continuous on its domain $\mathbb{R}$, we deduce that f is continuous on the bounded closed interval $[-2, 2]$. Thus f is uniformly continuous on $[-2, 2]$, by Theorem F19.

Unit F2
Differentiation

Introduction

In Unit D4 *Continuity* you studied *continuous* real functions and saw that they share some important properties; for example, they all satisfy the Intermediate Value Theorem, the Extreme Value Theorem and the Boundedness Theorem. However, many of the most interesting properties of functions are obtained only when we further restrict our attention to *differentiable* functions.

You have already met the idea of *differentiating* a given real function f; that is, finding the gradient of the tangent to the graph $y = f(x)$ at those points of the graph where a tangent exists. The gradient of the tangent at the point $(c, f(c))$ is called the *derivative* of f at c, and is written as $f'(c)$. In this unit we investigate which real functions are differentiable, and we discuss some of the important properties that all differentiable functions possess.

Many of the techniques of differentiation and the properties of differentiable functions covered in this unit will not be new to you, but our main focus here is on establishing a rigorous foundation for these ideas. We will make frequent use of results on limits and continuity that you met in previous analysis units.

1 Differentiable functions

In this section we define what it means for a real function f to be differentiable at a point c, and we establish that certain basic functions are differentiable and find their derivatives. We also consider functions which possess higher derivatives, that is, functions which can be differentiated more than once. Finally, we show that differentiable functions are continuous, and also prove that the blancmange function, which was shown to be continuous in Unit F1 *Limits*, is in fact not differentiable at any point in its domain.

1.1 What is differentiability?

Differentiability arises from the geometric concept of the *tangent* to a graph. The tangent to the graph $y = f(x)$ at the point $(c, f(c))$, if it exists, is the line through the point $(c, f(c))$ whose direction is the 'limiting direction' of the chords joining the points $(c, f(c))$ and $(x, f(x))$ as x tends to c. This idea is illustrated in Figure 1.

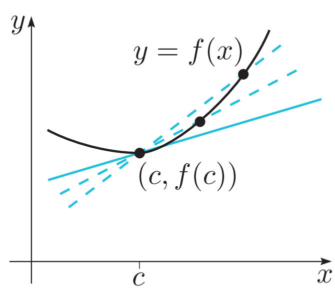
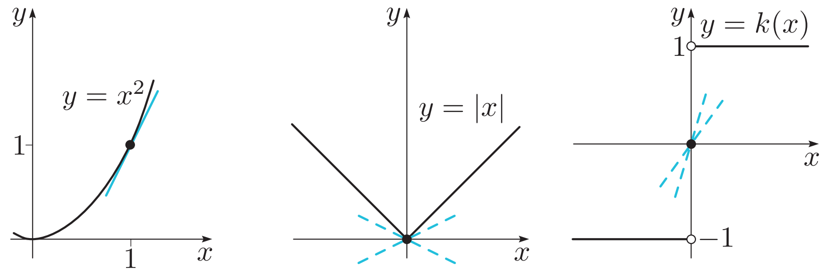


Figure 1 Constructing the tangent to a graph

The three examples shown in Figure 2 below illustrate some of the possibilities that can occur when we try to construct tangents in particular instances.



(a) (b) (c)

Figure 2 Different possibilities that occur when trying to construct tangents (the rule for the function k is given in the text)

The function

$$f(x) = x^2 \quad (x \in \mathbb{R})$$

is continuous on $\mathbb{R}$, and its graph (Figure 2(a)) has a tangent at each point; for example, the line $y = 2x - 1$ is the tangent to the graph at the point $(1, 1)$.

The function

$$g(x) = |x| \quad (x \in \mathbb{R})$$

is also continuous on $\mathbb{R}$, but its graph (Figure 2(b)) does not have a tangent at the point $(0, 0)$: no line through the point $(0, 0)$ is a tangent to the graph. This is because the 'limiting direction' of the chords described earlier is different depending on whether $(0, 0)$ is approached from the left or from the right. However, there is a tangent at every other point of the graph.

Finally, the function

$$k(x) = \begin{cases} -1, & x < 0, \\ 0, & x = 0, \\ 1, & x > 0, \end{cases}$$

is discontinuous at 0, and no line through the point $(0, 0)$ is a tangent to the graph (see Figure 2(c)). However, there is a tangent at every other point of the graph. We now make these ideas precise, by using the concept of limit (which you studied in Unit F1) to pin down what we mean by 'limiting direction'. We define the **gradient** (or slope) of the graph at $(c, f(c))$ to be the limit, as x tends to c, of the gradient of the chord through the points $(c, f(c))$ and $(x, f(x))$. The gradient of this chord is

$$\frac{f(x) - f(c)}{x - c}, \quad \text{where } x \neq c,$$

as illustrated in Figure 3. This expression is called the **difference quotient** for f at c. Thus the gradient of the graph of f at the point $(c, f(c))$ is

$$\lim_{x\to c} \frac{f(x)-f(c)}{x-c}, \qquad (1)$$

provided that this limit exists. We call this limit the *derivative* of f at c.

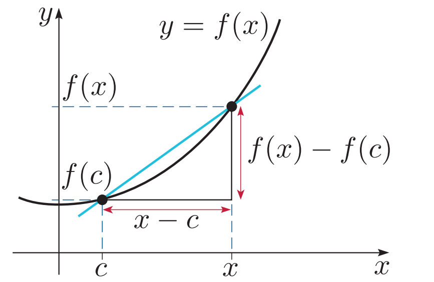


Figure 3 The chord joining $(c, f(c))$ and $(x, f(x))$

Sometimes it is more convenient to use an equivalent form of the difference quotient. If we replace x by $c+h$, then '$x \to c$' in expression (1) is equivalent to '$h \to 0$'. The difference quotient for f at c is then

$$Q(h) = \frac{f(c+h)-f(c)}{h}, \quad \text{where } h \neq 0,$$

as illustrated in Figure 4, and the gradient of the graph of f at $(c, f(c))$, that is, the derivative of f at c, is given by

$$\lim_{h\to 0} Q(h),$$

provided that this limit exists.

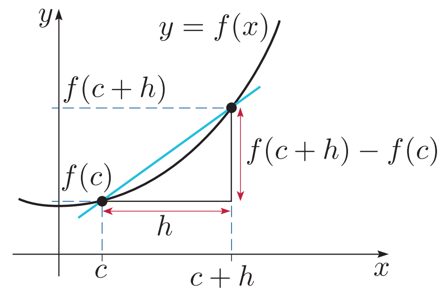


Figure 4 The chord joining $(c, f(c))$ and $(c+h, f(c+h))$

To formalise this concept, we need to ensure that f is defined near the point c, so we assume that c lies in some open interval I in the domain of f.

Definitions

Let f be defined on an open interval I, and let $c \in I$. Then the **derivative** of f at c is

$$\lim_{x\to c} \frac{f(x)-f(c)}{x-c},$$

that is,

$$\lim_{h\to 0} Q(h), \quad \text{where} \quad Q(h) = \frac{f(c+h)-f(c)}{h},$$

provided that this limit exists. If the limit exists, then we say that f is **differentiable at c**. If f is differentiable at each point of its domain, then we say that f is **differentiable** (on its domain).

The derivative of f at c is denoted by $f'(c)$ and the function $f' : x \longmapsto f'(x)$ is called the **derivative**, or sometimes the **derived function**, of f.

The operation of obtaining $f'(x)$ from $f(x)$ is called **differentiation**.

Remarks

1. The word 'differentiable' arises because the definition involves the *differences* $f(x) - f(c)$ and $x - c$.
2. The above box defines the notation for derivatives that we will normally use in this module, but there are several others in common use. Sometimes f' is denoted by Df and $f'(x)$ is denoted by $Df(x)$. Another frequently used notation is Leibniz notation, in which $f'(x)$ is written as $\dfrac{dy}{dx}$, where $y = f(x)$.
3. Note that we require f to be defined on an open interval containing c because, by the definition of the limit, x can approach c along any sequence of points in a punctured neighbourhood of c. In the next subsection we will define one-sided derivatives using the corresponding one-sided limits that you met in Unit F1.
4. The existence of the derivative $f'(c)$ is not quite equivalent to the existence of a tangent to the graph $y = f(x)$ at the point $(c, f(c))$. If $\lim\limits_{x\to c} \dfrac{f(x) - f(c)}{x - c}$ exists, then the graph certainly has a tangent at the point $(c, f(c))$, and the gradient of the tangent is the value of the limit. However, the converse is not necessarily true. The graph may have a *vertical* tangent at the point $(c, f(c))$, in which case

$$\frac{f(x) - f(c)}{x - c} \to \infty \text{ (or } -\infty) \quad \text{as } x \to c.$$

In this case, $\lim\limits_{x\to c} \dfrac{f(x) - f(c)}{x - c}$ does not exist, so f is not differentiable at c.

The idea of finding the gradient of a graph which is not a straight line was one of the first steps in the development of calculus. The major figures in this development were Sir Isaac Newton (1642–1727) in England and Gottfried Wilhelm Leibniz (1646–1716) in Germany, to whom we owe the 'dy/dx' notation and the names 'differential calculus' and 'integral calculus'. The names 'derived function' and 'derivative' and the notation f' were introduced much later by Lagrange (1736–1813). Another mathematician who played a role in the early development of calculus was Pierre de Fermat (1601?–1665) in France.

Worked Exercise F9

Prove that the function $f(x) = x^3$ is differentiable at each point $c \in \mathbb{R}$, and determine $f'(c)$.

Solution

To prove that a function is differentiable, it is usually most convenient to use the $Q(h)$ form of the difference quotient.

The difference quotient for f at c is

$$\begin{aligned} Q(h) = \frac{f(c+h) - f(c)}{h} &= \frac{(c+h)^3 - c^3}{h} \\ &= \frac{(c^3 + 3c^2h + 3ch^2 + h^3) - c^3}{h} \\ &= 3c^2 + 3ch + h^2, \quad \text{where } h \neq 0. \end{aligned}$$

Thus $Q(h) \to 3c^2$ as $h \to 0$, so f is differentiable at c, with $f'(c) = 3c^2$.

Here we have used the Combination Rules for limits without mentioning them explicitly.

We use c to denote a particular point where we are testing for differentiability. However, when stating the rule of a derivative, we replace c by the usual variable x. Thus the derivative of f in Worked Exercise F9 is $f'(x) = 3x^2$.

To prove from the definition that a function is *not* differentiable at a point, we need to show that the limit of the difference quotient does not exist. The following strategy is based on Strategy F1 from Unit F1.

Strategy F5

To prove that a function is not differentiable at a point, show that $\lim\limits_{h\to 0} Q(h)$ does not exist by doing either of the following.

- Find two null sequences (h_n) and (k_n) with non-zero terms such that the sequences $(Q(h_n))$ and $(Q(k_n))$ have different limits.
- Find a null sequence (h_n) with non-zero terms such that $Q(h_n) \to \infty$ or $Q(h_n) \to -\infty$.

The next worked exercise illustrates how to apply this strategy.

Worked Exercise F10

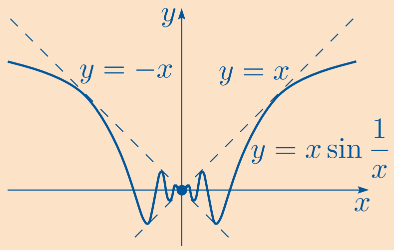
Prove that the function

$$f(x) = \begin{cases} x \sin \dfrac{1}{x}, & x \neq 0, \\ 0, & x = 0, \end{cases}$$

is not differentiable at 0.

Solution

The graph of f is shown below.



The difference quotient for f at $c = 0$ is

$$\begin{aligned} Q(h) &= \frac{f(h) - f(0)}{h} \\ &= \frac{h \sin(1/h) - 0}{h} = \sin(1/h), \quad \text{where } h \neq 0. \end{aligned}$$

Since $\sin(1/h)$ oscillates infinitely often in any interval containing the origin, giving different tangents at different points, we use the first method of Strategy F5.

Consider the two sequences

$$h_n = \frac{1}{n\pi} \quad \text{and} \quad k_n = \frac{1}{\left(2n + \frac{1}{2}\right)\pi}, \quad n = 1, 2, \ldots.$$

Then (h_n) and (k_n) are null sequences with non-zero terms, chosen so that

$$Q(h_n) = \sin(1/h_n) = \sin(n\pi) = 0 \to 0 \text{ as } n \to \infty$$

and

$$Q(k_n) = \sin(1/k_n) = \sin\left(2n + \tfrac{1}{2}\right)\pi = 1 \to 1 \text{ as } n \to \infty.$$

Since these limits are different, f is not differentiable at 0.

Other choices of (h_n) and (k_n) are possible; for example,

$$h_n = \frac{1}{2n\pi} \text{ and } k_n = \frac{1}{\left(2n - \frac{1}{2}\right)\pi}, \quad n = 1, 2, \ldots,$$

giving $Q(h_n) = 0$ and $Q(k_n) = -1$.

Worked Exercise F10 shows that the domain of a derivative f' can be smaller than the domain of f.

Exercise F15

(a) Prove that the function $f(x) = 1/x$ is differentiable at each point $c \in \mathbb{R} - \{0\}$, and determine $f'(c)$.

(b) Prove that the function

$$f(x) = \begin{cases} x^2 \cos(1/x), & x \neq 0, \\ 0, & x = 0, \end{cases}$$

is differentiable at 0, and determine $f'(0)$.

Hint: Use the Squeeze Rule for limits from Subsection 1.3 of Unit F1.

(c) Prove that the function $f(x) = |x|$ is not differentiable at 0.

(d) Prove that the function

$$f(x) = \begin{cases} |x|^{1/2} \sin(1/x), & x \neq 0, \\ 0, & x = 0, \end{cases}$$

is not differentiable at 0.

We now establish the derivatives of some basic functions.

Theorem F21 Basic derivatives

(a) If $f(x) = k$, where $k \in \mathbb{R}$, then $f'(x) = 0$.

(b) If $f(x) = x^n$, where $n \in \mathbb{N}$, then $f'(x) = nx^{n-1}$.

(c) If $f(x) = \sin x$, then $f'(x) = \cos x$.

(d) If $f(x) = \cos x$, then $f'(x) = -\sin x$.

(e) If $f(x) = e^x$, then $f'(x) = e^x$.

Proof (a) If $f(x) = k$, then the difference quotient for f at any point is 0, so $f'(x) = 0$.

(b) The difference quotient for f at c is

$$Q(h) = \frac{f(c+h) - f(c)}{h} = \frac{(c+h)^n - c^n}{h}.$$

Recall that, for $a, b \in \mathbb{R}$ and $n \in \mathbb{N}$,

$$(a+b)^n = a^n + na^{n-1}b + \frac{n(n-1)}{2!}a^{n-2}b^2 + \cdots + b^n.$$

This is the Binomial Theorem, which you met in Subsection 3.4 of Unit D1 *Numbers*.

So, by the Binomial Theorem,

$$\begin{aligned} Q(h) &= \frac{1}{h}\left(c^n + nc^{n-1}h + \frac{n(n-1)}{2}c^{n-2}h^2 + \cdots + h^n - c^n\right) \\ &= nc^{n-1} + \frac{n(n-1)}{2}c^{n-2}h + \cdots + h^{n-1}, \quad \text{where } h \neq 0. \end{aligned}$$

Thus $Q(h) \to nc^{n-1}$ as $h \to 0$, so f is differentiable at c for any $c \in \mathbb{R}$, and $f'(x) = nx^{n-1}$.

(c) The difference quotient for f at c is

$$\begin{aligned} Q(h) &= \frac{\sin(c+h) - \sin c}{h} \\ &= \frac{\sin c \cos h + \cos c \sin h - \sin c}{h} \\ &= \cos c\left(\frac{\sin h}{h}\right) + \sin c\left(\frac{\cos h - 1}{h}\right), \quad \text{where } h \neq 0. \end{aligned}$$

You met the limits

$$\lim_{h\to 0} \frac{\sin h}{h} = 1 \quad \text{and} \quad \lim_{h\to 0} \frac{\cos h - 1}{h} = 0$$

in Subsections 1.1 and 1.3 of Unit F1, respectively.

Thus

$$Q(h) \to \cos c \times 1 + \sin c \times 0 = \cos c \;\text{ as } h \to 0,$$

so $f'(x) = \cos x$, as required.

(d) The proof of this part is similar to that of part (c). We omit the details – you may like to write out the proof for yourself.

(e) The difference quotient for f at c is

$$\begin{aligned} Q(h) &= \frac{e^{c+h} - e^c}{h} \\ &= e^c\left(\frac{e^h - 1}{h}\right), \quad \text{where } h \neq 0. \end{aligned}$$

You saw that $\lim_{h\to 0} \frac{e^h - 1}{h} = 1$ in Subsection 1.3 of Unit F1.

Thus

$$Q(h) \to e^c \times 1 = e^c \;\text{ as } h \to 0,$$

so $f'(x) = e^x$, as required. ■

In general, when we differentiate a function f, we obtain a new function f' whose domain may be smaller than that of f. The notion of differentiability can then be applied to the function f', yielding another function whose domain consists of those points where f' is differentiable.

Definitions

Let f be differentiable on an open interval I, and let $c \in I$. If the derivative f' is differentiable at c, then we say that f is **twice differentiable at c**, and the number $f''(c) = (f')'(c)$ is called the **second derivative of f at c**. The function f'', also denoted by $f^{(2)}$, is called the **second derivative** (or **second derived function**) of f.

Similarly, we can define the **higher-order derivatives** of f, denoted by $f^{(3)} = f'''$, $f^{(4)}$, and so on.

Remarks

1. You may meet many different ways of denoting the second derivative of a function. For example, f'' is sometimes denoted by Df' or $D^2(f)$.

   In Leibniz notation $f''(x)$ is written as $\dfrac{d^2y}{dx^2}$, where $y = f(x)$.
2. Some functions can be differentiated as many times as we like. For example, if $f(x) = e^x$, then

   $$f'(x) = e^x, \quad f''(x) = e^x, \quad f^{(3)}(x) = e^x, \quad \ldots.$$

   However, not every derivative is differentiable at all points of its domain; you will see an example of this in Exercise F16.

1.2 One-sided derivatives

Although we know that the function $f(x) = |x|$ is not differentiable at 0, the graph of f shown in Figure 5 suggests that chords which join the origin $(0, 0)$ to points $(h, f(h))$ have gradients equal to 1 if $h > 0$, and equal to -1 if $h < 0$. This example suggests the concept of a *one-sided derivative*.

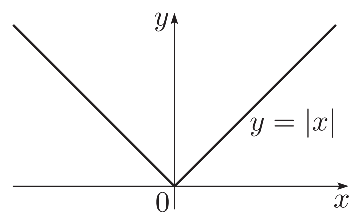


Figure 5 The graph of $f(x) = |x|$

Definitions

Let f be defined on an interval I, and let $c \in I$. Then the **left derivative** of f at c is

$$f'_L(c) = \lim_{x \to c^-} \frac{f(x) - f(c)}{x - c} = \lim_{h \to 0^-} Q(h),$$

provided that this limit exists. If the limit exists, we say that f is **left differentiable** at c.

Similarly, the **right derivative** of f at c is

$$f'_R(c) = \lim_{x \to c^+} \frac{f(x) - f(c)}{x - c} = \lim_{h \to 0^+} Q(h),$$

provided that this limit exists. If the limit exists, we say that f is **right differentiable** at c.

Remarks

1. The definition uses the concept of right and left limits which were defined in Subsection 1.4 of Unit F1.
2. In some texts, a function f defined on a bounded closed interval I is *defined* to be differentiable on I if f has a derivative at each interior point of I and appropriate one-sided derivatives at each endpoint of I.

The next theorem establishes the relationship between derivatives and one-sided derivatives.

Theorem F22

Let f be defined on an open interval I, and let $c \in I$.

(a) If f is differentiable at c, then f is both left differentiable and right differentiable at c, and

$$f'_L(c) = f'_R(c) = f'(c).$$

(b) If f is both left differentiable and right differentiable at c, and $f'_L(c) = f'_R(c)$, then f is differentiable at c and

$$f'(c) = f'_L(c) = f'_R(c).$$

Theorem F22 is closely related to the next theorem, the Glue Rule for differentiable functions, which is illustrated in Figure 6 and stated below. We omit the proofs of both these results as the ideas involved are very similar to those used in the proof of the Glue Rule for continuous functions (see Subsection 2.2 of Unit D4).

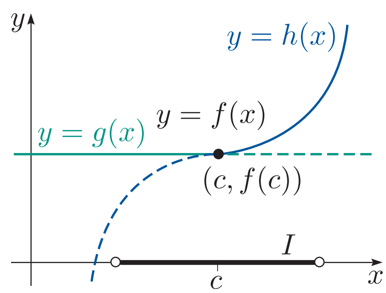


Figure 6 The Glue Rule for differentiable functions

Theorem F23 Glue Rule for differentiable functions

Let f be defined on an open interval I, and let $c \in I$. If there are functions g and h defined on I such that

1. $f(x) = g(x)$, for $x \in I$, $x < c$,
   $f(x) = h(x)$, for $x \in I$, $x > c$,
2. $f(c) = g(c) = h(c)$, and
3. g and h are differentiable at c,

then f is differentiable at c if and only if $g'(c) = h'(c)$. If f is differentiable at c, then $f'(c) = g'(c) = h'(c)$.

The Glue Rule enables us to determine whether or not certain hybrid functions are differentiable at particular points, without using the definition of differentiability. The next worked exercise illustrates how to use the Glue Rule in this way.

Worked Exercise F11

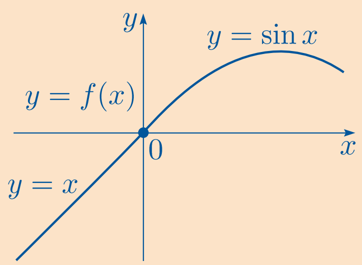
Use the Glue Rule to prove that the function

$$f(x) = \begin{cases} x, & x < 0, \\ \sin x, & x \geq 0, \end{cases}$$

is differentiable at 0, and determine $f'(0)$.

Solution

The graph of f is shown below.



Let I be the open interval $\mathbb{R}$ and define

$$g(x) = x \quad (x \in \mathbb{R}) \quad \text{and} \quad h(x) = \sin x \quad (x \in \mathbb{R}).$$

Then f is defined on I and $0 \in I$. Also,

$$\begin{aligned} f(x) &= g(x), \quad \text{for } x < 0, \\ f(x) &= h(x), \quad \text{for } x > 0, \end{aligned}$$

so condition 1 of the Glue Rule holds with $c = 0$.

Furthermore, $f(0) = g(0) = h(0) = 0$, so condition 2 holds, and g and h are differentiable with

$$g'(x) = 1 \quad \text{and} \quad h'(x) = \cos x,$$

so condition 3 holds.

Since $g'(0) = 1 = h'(0)$, we deduce that f is differentiable at 0, with $f'(0) = 1$, by the Glue Rule.

If we want to prove that the function f in Worked Exercise F11 is differentiable at a point c other than 0, then we can use the fact that differentiability is a *local property* of the function; that is, it depends on the behaviour of the function in any open interval (no matter how short) containing c. Thus

$$f'(x) = \begin{cases} g'(x) = 1, & x < 0, \\ h'(x) = \cos x, & x > 0, \end{cases}$$

and hence, on combining this result with Worked Exercise F11, we have

$$f'(x) = \begin{cases} 1, & x \leq 0, \\ \cos x, & x > 0. \end{cases}$$

You can use this approach in the following exercise.

Exercise F16

Prove that the function

$$f(x) = \begin{cases} -x^2, & x < 0, \\ x^2, & x \geq 0, \end{cases}$$

is differentiable, and has derivative $f'(x) = 2|x|$.

Since the function $f'(x) = 2|x|$ is not differentiable at 0, Exercise F16 shows that a derivative need not be differentiable at all points of its domain.

Another consequence of the fact that differentiability is a local property is that the *restriction* of a differentiable function to an open subinterval of its domain gives a new differentiable function. For example, the function

$$f(x) = x^2 \quad (x \in (2,3))$$

is differentiable, since $f(x) = x^2$ is differentiable on $\mathbb{R}$.

1.3 Continuity and differentiability

Next we discuss the relationship between continuity and differentiability. First we show that a differentiable function is continuous.

Theorem F24

Let f be defined on an open interval I, and let $c \in I$. If f is differentiable at c, then f is continuous at c.

Proof If f is differentiable at c, then

$$\lim_{x \to c} \frac{f(x) - f(c)}{x - c} = f'(c).$$

For $x \in I$ and $x \neq c$, we have

$$f(x) - f(c) = \frac{f(x) - f(c)}{(x - c)} \times (x - c).$$

Hence, by the Combination Rules for limits (see Subsection 1.3 of Unit F1),

$$f(x) - f(c) \to f'(c) \times 0 = 0 \text{ as } x \to c.$$

Thus $f(x) \to f(c)$ as $x \to c$, so f is continuous at c. ■

The following corollary gives us a test for non-differentiability; it is simply the contrapositive of Theorem F24 (and so is equivalent to it).

Corollary F25

Let f be defined on an open interval I, and let $c \in I$. If f is discontinuous at c, then f is not differentiable at c.

For example, the function

$$k(x) = \begin{cases} -1, & x < 0, \\ 0, & x = 0, \\ 1, & x > 0, \end{cases}$$

is discontinuous at 0 because $\lim_{x \to 0^+} k(x) = 1 \neq k(0)$. Thus k is not differentiable at 0, by Corollary F25.

It is important to remember that a function can be continuous at a point but *not* differentiable at that point; for example, the modulus function is continuous at all points of $\mathbb{R}$, but it is not differentiable at 0, as you saw in Exercise F15(c). This example can readily be modified to produce a continuous function which is not differentiable at any given finite set of points. It is even possible for a function to be continuous throughout its domain but *nowhere* differentiable, as we discuss below.

A continuous nowhere-differentiable function (optional)

The remainder of this section is not assessed, and is included only for your interest.

In the nineteenth century, when the concepts of continuity and differentiability were first made precise, it was widely believed that if a function is continuous at all points of an interval, then it must be differentiable at most points of that interval. However, it turns out that there exist functions which are continuous everywhere but differentiable *nowhere*. The first example was found as early as 1834 by Bernard Bolzano (1781–1848), but his pioneering work on analysis was not widely known. The first well-known example was constructed by Karl Weierstrass (1815–1897) in 1872. Such 'pathological' functions were regarded by some with suspicion. For example, the French mathematician Charles Hermite (1822–1901) wrote to a colleague in 1893, 'I recoil in fear and loathing from that deplorable evil: continuous functions with no derivatives.' However, in modern times it has been shown that such functions are in some sense normal and even useful.

In Subsection 3.2 of Unit F1 you met the *blancmange function* and saw that it is continuous at all points in its domain $\mathbb{R}$. We now prove that this function is nowhere differentiable. Recall that the blancmange function B is defined as follows:

$$B(x) = \sum_{n=0}^{\infty} \frac{1}{2^n} s(2^n x) = s(x) + \tfrac{1}{2}s(2x) + \tfrac{1}{4}s(4x) + \cdots \quad (x \in \mathbb{R}), \tag{2}$$

where s is the sawtooth function

$$s(x) = \begin{cases} x - \lfloor x \rfloor, & 0 \le x - \lfloor x \rfloor \le \frac{1}{2}, \\ 1 - (x - \lfloor x \rfloor), & \frac{1}{2} < x - \lfloor x \rfloor < 1. \end{cases}$$

The graph of the function s has a 'corner' at each point of the form $k/2$, where $k \in \mathbb{Z}$, so the graph of each function $x \longmapsto 2^{-n}s(2^n x)$ has a corner at each point of the form $k/2^{n+1}$, where $k \in \mathbb{Z}$. This suggests that the graph of the function B, shown in Figure 7, does not have a tangent at any point of $\mathbb{R}$, and we now prove that this is the case.

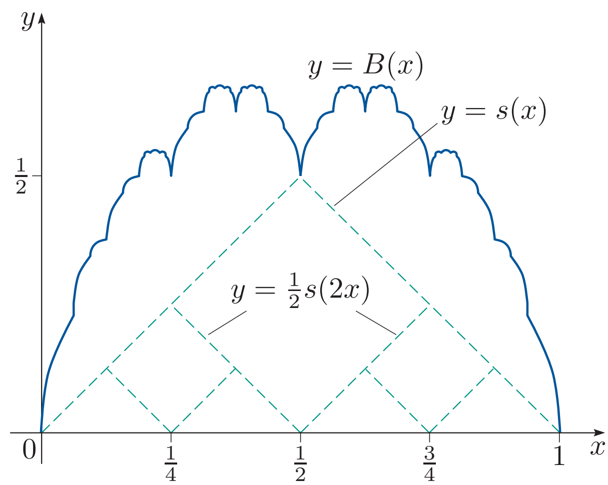


Figure 7 The blancmange function

The name 'blancmange function' was used in the 1980s by the English mathematician David Tall, who remarked that B is nowhere differentiable because it wobbles too much!

Theorem F26

The blancmange function is not differentiable at any point $c \in \mathbb{R}$.

To prove Theorem F26, we first prove three preliminary lemmas.

As we will need to refer to the difference quotients of several functions, we adopt the following notation which shows the dependence on both the point c and the function f:

$$Q_{c,f}(h) = \frac{f(c+h) - f(c)}{h}, \quad \text{where } h \neq 0.$$

Lemma F27

Let B be the blancmange function. Then for $m = 1, 2, \ldots,$ we have

$$B(x) = s(x) + \frac{1}{2}s(2x) + \cdots + \frac{1}{2^{m-1}}s(2^{m-1}x) + \frac{1}{2^m}B(2^m x),$$

and the function

$$x \longmapsto s(x) + \frac{1}{2}s(2x) + \cdots + \frac{1}{2^{m-1}}s(2^{m-1}x)$$

is linear on all intervals of the form $[p\,2^{-m}, (p+1)2^{-m}]$, where $p \in \mathbb{Z}$.

Proof Recall that a function is *linear* if it has a rule of the form $x \longmapsto ax + b$ for some $a, b \in \mathbb{R}$.

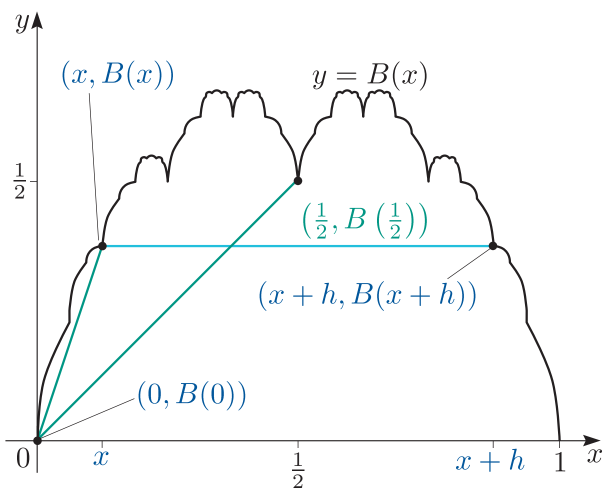
This follows directly from equation (2) and the definition of s. ■

Lemma F28

Let B be the blancmange function. Then for each $x \in [0, 1]$, there exist $h, k \neq 0$ such that

$$x + h,\ x + k \in [0, 1] \quad \text{and} \quad |Q_{x,B}(h) - Q_{x,B}(k)| \geq 1.$$

Proof The proof is illustrated in the following figure.



Observe that the graph of $y = B(x)$ is symmetrical about the line $x = \frac{1}{2}$, and that $B(x) \geq x$ for $x \in \left[0, \frac{1}{2}\right]$.

We can assume by symmetry that $0 \le x \le \frac{1}{2}$. We now choose h so that $x + h = 1 - x$, and k so that $x + k = 0$ (or $k = \frac{1}{2}$ if $x = 0$). Then

$$Q_{x,B}(h) = \frac{B(x+h) - B(x)}{h} = \frac{B(1-x) - B(x)}{h} = 0,$$

by symmetry. Also, for $x \neq 0$,

$$Q_{x,B}(k) = \frac{B(x+k) - B(x)}{k} = \frac{B(0) - B(x)}{-x} \ge 1,$$

because $B(0) = 0$ and $B(x) \ge x$ for $x \in \left[0, \frac{1}{2}\right]$. A similar argument shows that the same result holds for $x = 0$. Thus the inequality in the statement of the lemma follows, as required. ■

Lemma F29

Given any real function f and a linear function $g(x) = ax + b$, the corresponding difference quotients of the functions f and $f + g$ always differ by a, the gradient of the linear function g.

Proof This holds because

$$Q_{c,f+g}(h) - Q_{c,f}(h) = Q_{c,g}(h) = (g(c+h) - g(c))/h = a,$$

for any $c \in \mathbb{R}$ and $h \neq 0$. ■

We now use these three lemmas to prove that the blancmange function is nowhere differentiable.

Proof of Theorem F26 Let $c \in \mathbb{R}$ and choose integers p_m, $m = 0, 1, 2, \ldots$, such that $c \in I_m$, where

$$I_m = \left[p_m 2^{-m},\ (p_m + 1)2^{-m}\right].$$

Note that (I_m) is a 'nested' sequence of closed intervals, each of which contains the point c.

Now it follows from Lemma F27 that, on the interval I_m, the function B is the sum of a linear function and the function $B_m(x) = 2^{-m}B(2^m x)$, which is obtained from B by scaling in both the x- and y-directions by the factor 2^{-m}.

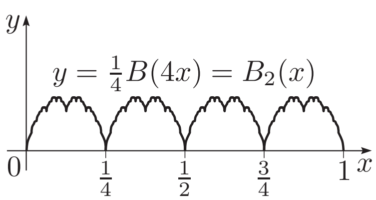


Figure 8 The graph of $B_2(x)$

The graph of B_m on the interval I_m is a scaled-down copy of the graph of B on $[0, 1]$ (see Figure 8): it is a mini-blancmange!

Difference quotients are unchanged by such a scaling, so Lemma F28 implies that there exist $h_m, k_m \neq 0$ such that

$$c + h_m,\ c + k_m \in I_m \quad \text{and} \quad |Q_{c,B_m}(h_m) - Q_{c,B_m}(k_m)| \ge 1.$$

Therefore, by Lemma F29,

$$|Q_{c,B}(h_m) - Q_{c,B}(k_m)| = |Q_{c,B_m}(h_m) - Q_{c,B_m}(k_m)| \geq 1. \tag{3}$$

This holds because the difference quotients of the linear part of the function B cancel out.

Now since c, $c + h_m$, $c + k_m \in I_m$, for $m = 1, 2, \ldots$, and I_m has length 2^{-m}, we have $h_m \to 0$ and $k_m \to 0$. Thus if B is differentiable at c, then it follows from the definition of differentiability that

$$Q_{c,B}(h_m) \to B'(c) \quad \text{and} \quad Q_{c,B}(k_m) \to B'(c) \;\text{ as } m \to \infty.$$

But this contradicts inequality (3), so B is not differentiable at c. ■

Finally we note that, since the function B is continuous, it follows from the Extreme Value Theorem (see Subsection 3.3 of Unit D4) that it must have a maximum and a minimum on the interval $[0, 1]$. The minimum is $B(0) = B(1) = 0$, but the maximum is not so clear. In fact, it can be shown that the maximum is $B\left(\frac{1}{3}\right) = \frac{2}{3}$, and that this value is taken at infinitely many points in $[0, 1]$; a strange function indeed!

2 Rules for differentiation

In Section 1 you saw how to show that various basic functions are differentiable on $\mathbb{R}$, by working directly with the definition of a differentiable function. In this section you will see that we can often show that a function is differentiable by using the Combination Rules, the Composition Rule (also called the Chain Rule) and the Inverse Function Rule for differentiable functions.

We will use these rules to determine the derivatives of many more functions. The most important of these (together with the basic functions from Section 1) are summarised in a table of standard derivatives, which can be found at the end of this unit and in the module Handbook.

2.1 Combination Rules

The Combination Rules for differentiable functions are a consequence of the Combination Rules for limits that you met in Subsection 1.3 of Unit F1.

Theorem F30 Combination Rules for differentiable functions

Let f and g be defined on an open interval I, and let $c \in I$. If f and g are differentiable at c, then so are the following functions.

Sum Rule $f+g$, with derivative

$$(f+g)'(c) = f'(c) + g'(c)$$

Multiple Rule λf, for $\lambda \in \mathbb{R}$, with derivative

$$(\lambda f)'(c) = \lambda f'(c)$$

Product Rule fg, with derivative

$$(fg)'(c) = f'(c)g(c) + f(c)g'(c)$$

Quotient Rule f/g, provided that $g(c) \neq 0$, with derivative

$$\left(\frac{f}{g}\right)'(c) = \frac{g(c)f'(c) - f(c)g'(c)}{(g(c))^2}.$$

Each of these rules can be expressed in any of the alternative notations commonly used for derivatives. For example, in Leibniz notation, the Product Rule becomes

$$\text{if } y = uv, \text{ then } \frac{dy}{dx} = \frac{du}{dx}v + u\frac{dv}{dx}.$$

Similarly, the Quotient Rule is written

$$\text{if } y = u/v, \text{ then } \frac{dy}{dx} = \frac{1}{v^2}\left(v\frac{du}{dx} - u\frac{dv}{dx}\right).$$

Proof of Theorem F30 To prove each rule, we denote the function that we wish to show is differentiable by F, and express the difference quotient for F in terms of the difference quotients for f and g.

Suppose that f and g are both differentiable at c.

Sum Rule Let $F = f + g$. Then

$$\begin{aligned}\frac{F(x) - F(c)}{x - c} &= \frac{(f(x) + g(x)) - (f(c) + g(c))}{x - c} \\ &= \frac{f(x) - f(c)}{x - c} + \frac{g(x) - g(c)}{x - c} \\ &\to f'(c) + g'(c) \quad \text{as } x \to c,\end{aligned}$$

by the Sum Rule for limits, since f and g are differentiable at c. Thus F is differentiable at c, with derivative

$$F'(c) = f'(c) + g'(c).$$

Multiple Rule This is a special case of the Product Rule, with $g(x) = \lambda$.

Product Rule Let $F = fg$. Then

$$\begin{aligned}\frac{F(x)-F(c)}{x-c} &= \frac{f(x)g(x)-f(c)g(c)}{x-c} \\ &= \frac{f(x)-f(c)}{x-c}\,g(x) + f(c)\,\frac{g(x)-g(c)}{x-c} \\ &\to f'(c)g(c)+f(c)g'(c) \quad \text{as } x \to c,\end{aligned}$$

by the Combination Rules for limits, since f and g are differentiable at c, and g is continuous at c by Theorem F24, so that $g(x) \to g(c)$ as $x \to c$.

Thus F is differentiable at c, with derivative

$$F'(c) = f'(c)g(c) + f(c)g'(c).$$

Quotient Rule We first use the ε-δ definition of continuity that you met in Subsection 3.1 of Unit F1 to show that, since $g(c) \neq 0$ and g is continuous at c, then g must be non-zero on an open interval containing c. To show this, we take $\varepsilon = \frac{1}{2}|g(c)|$ in the definition.

Let $F = f/g$. Since g is continuous at c and $g(c) \neq 0$, there exists $\delta > 0$ such that $J = (c-\delta, c+\delta) \subseteq I$ and

$$|g(x)-g(c)| < \tfrac{1}{2}|g(c)|, \quad \text{for all } x \text{ with } |x-c| < \delta.$$

In particular, this shows that $g(x) \neq 0$ for $x \in J$, so the domain of F contains J. Then, for $x \in J$,

$$\begin{aligned}\frac{F(x)-F(c)}{x-c} &= \frac{1}{x-c}\left(\frac{f(x)}{g(x)} - \frac{f(c)}{g(c)}\right) \\ &= \frac{f(x)g(c)-f(c)g(x)}{(x-c)g(x)g(c)} \\ &= \frac{g(c)(f(x)-f(c)) - f(c)(g(x)-g(c))}{(x-c)g(x)g(c)} \\ &= \frac{1}{g(x)g(c)}\left(g(c)\,\frac{f(x)-f(c)}{x-c} - f(c)\,\frac{g(x)-g(c)}{x-c}\right) \\ &\to \frac{g(c)f'(c)-f(c)g'(c)}{(g(c))^2} \quad \text{as } x \to c,\end{aligned}$$

by the Combination Rules for limits, since f and g are differentiable at c, and g is continuous at c.

Thus F is differentiable at c, with derivative

$$F'(c) = \frac{g(c)f'(c)-f(c)g'(c)}{(g(c))^2}.$$

■

Since, by Theorem F21(b), the derivative of $f(x) = x^n$ is $f'(x) = nx^{n-1}$ for any $n \in \mathbb{N}$, it follows from the Combination Rules that any polynomial function is differentiable on $\mathbb{R}$ and that its derivative can be obtained by differentiating the polynomial term by term. We state this result as a corollary of the Combination Rules.

Corollary F31

Let

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \quad (x \in \mathbb{R}),$$

where $a_0, a_1, \ldots, a_n \in \mathbb{R}$. Then p is differentiable on $\mathbb{R}$, with derivative

$$p'(x) = a_1 + 2a_2x + \cdots + na_nx^{n-1} \quad (x \in \mathbb{R}).$$

Furthermore, since a rational function is a quotient of two polynomials, it follows from Corollary F31 and from the Quotient Rule that a rational function is differentiable at all points where its denominator is non-zero, that is, at all points of the domain of the function. Here is an example.

Worked Exercise F12

Prove that the function

$$f(x) = \frac{x^3}{x^2-1} \quad (x \in \mathbb{R} - \{-1, 1\})$$

is differentiable on its domain, and find its derivative.

Solution

The function f is a rational function of the form $f = p/q$, where $p(x) = x^3$ and $q(x) = x^2 - 1$, whose denominator q is non-zero on $\mathbb{R} - \{-1, 1\}$. Thus, by the Quotient Rule, f is differentiable on $\mathbb{R} - \{-1, 1\}$.

Here we write out all the steps in the application of the Quotient Rule, but it is not necessary for you to do this in your solutions.

Now $p'(x) = 3x^2$ and $q'(x) = 2x$. Thus, by the Quotient Rule, the derivative of f is

$$\begin{aligned} f'(x) &= \frac{(x^2-1)3x^2 - x^3(2x)}{(x^2-1)^2} \\ &= \frac{x^4 - 3x^2}{(x^2-1)^2} \quad (x \in \mathbb{R} - \{-1, 1\}). \end{aligned}$$

The Quotient Rule can also be used to show that the formula for differentiating $f(x) = x^n$, given in Theorem F21(b), remains valid if n is a *negative* integer.

We now give several exercises. It may not be necessary for you to work through all these exercises if you are confident that you can use the Combination Rules.

Exercise F17

Find the derivative of each of the following functions.

(a) $f(x) = x^7 - 2x^4 + 3x^3 - 5x + 1 \quad (x \in \mathbb{R})$

(b) $f(x) = \dfrac{x^2 + 1}{x^3 - 1} \quad (x \in \mathbb{R} - \{1\})$

(c) $f(x) = \sin x \cos x \quad (x \in \mathbb{R})$

(d) $f(x) = \dfrac{e^x}{3 + \sin x - 2\cos x} \quad (x \in \mathbb{R})$

Exercise F18

Find the third derivative of the function

$$f(x) = xe^x \quad (x \in \mathbb{R}).$$

In Section 1 we differentiated the functions sin, cos and exp. We now ask you to use these basic derivatives and the Combination Rules to find the derivatives of the other trigonometric functions and the three most common hyperbolic functions. The derivatives of these functions are included in the table of standard derivatives given at the end of this unit and in the module Handbook.

Exercise F19

Find the derivative of each of the following functions.

(a) $f(x) = \tan x \quad \left(x \neq (k + \frac{1}{2})\pi,\ k \in \mathbb{Z}\right)$

(b) $f(x) = \operatorname{cosec} x \quad (x \neq k\pi,\ k \in \mathbb{Z})$

(c) $f(x) = \sec x \quad \left(x \neq (k + \frac{1}{2})\pi,\ k \in \mathbb{Z}\right)$

(d) $f(x) = \cot x \quad (x \neq k\pi,\ k \in \mathbb{Z})$

Exercise F20

Find the derivative of each of the following functions.

(a) $f(x) = \sinh x \quad (x \in \mathbb{R})$

(b) $f(x) = \cosh x \quad (x \in \mathbb{R})$

(c) $f(x) = \tanh x \quad (x \in \mathbb{R})$

2.2 Composition Rule

In Subsection 2.1 we extended our stock of differentiable functions to include all polynomial, rational, trigonometric and hyperbolic functions. We also need to be able to differentiate functions such as

$$f(x) = \sin(\cos x) \quad (x \in \mathbb{R}),$$

which is the composite of the two differentiable functions sin and cos. To do this, we use the following Composition Rule for differentiable functions, which is commonly known as the Chain Rule.

Theorem F32 Composition Rule for differentiable functions (the Chain Rule)

Let f be defined on an open interval I, let g be defined on an open interval J such that $f(I) \subseteq J$, and let $c \in I$.

If f is differentiable at c and g is differentiable at $f(c)$, then $g \circ f$ is differentiable at c and

$$(g \circ f)'(c) = g'(f(c))f'(c).$$

Remarks

1. The Composition Rule tells us that 'a differentiable function of a differentiable function is differentiable', and gives us a formula for its derivative.
2. When written in Leibniz notation, the Composition Rule has a form that is easy to remember: if we put

   $$u = f(x) \quad \text{and} \quad y = g(u) = g(f(x)),$$

   then

   $$\frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dx}.$$
3. The Composition Rule can be extended to a composite of three or more functions; for example,

   $$(h \circ g \circ f)'(x) = h'(g(f(x)))g'(f(x))f'(x).$$

   In Leibniz notation, if we put

   $$v = f(x), \quad u = g(v) \quad \text{and} \quad y = h(u) = h(g(f(x))),$$

   then

   $$\frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dv} \times \frac{dv}{dx}.$$

   We often use this extended form of the Composition Rule without mentioning it explicitly.

Proof of Theorem F32 Let $F = g \circ f$. The difference quotient for F at c is

$$\frac{F(x) - F(c)}{x - c} = \frac{g(f(x)) - g(f(c))}{x - c}. \tag{4}$$

Let $y = f(x)$, where $x \in I$, and let $d = f(c)$. Then the right-hand side of equation (4) can be written

$$\left(\frac{g(y) - g(d)}{y - d}\right)\left(\frac{f(x) - f(c)}{x - c}\right), \quad \text{provided that } y \neq d. \tag{5}$$

To avoid the difficulty that expression (5) is undefined if $y = d$, which can occur in some situations, we introduce the function

$$h(y) = \begin{cases} \dfrac{g(y) - g(d)}{y - d}, & y \neq d, \\ g'(d), & y = d. \end{cases}$$

Since g is differentiable at d,

$$h(y) \to g'(d) \text{ as } y \to d;$$

and since $h(d) = g'(d)$, it follows that h is continuous at d.

By the Composition Rule for continuous functions (see Subsection 2.2 of Unit D4), and recalling that $y = f(x)$ and $d = f(c)$, we deduce that

$$(h \circ f)(x) = \begin{cases} \dfrac{g(f(x)) - g(f(c))}{f(x) - f(c)}, & f(x) \neq f(c), \\ g'(f(c)), & f(x) = f(c), \end{cases}$$

is continuous at c.

Next, note that if $f(x) \neq f(c)$, then equation (4) and expression (5) give

$$\frac{F(x) - F(c)}{x - c} = (h \circ f)(x)\left(\frac{f(x) - f(c)}{x - c}\right). \tag{6}$$

Equation (6) is also true when $f(x) = f(c)$, since both sides are then 0.

If we now let x tend to c in equation (6) and use the continuity at c of the function $h \circ f$, then we obtain

$$\frac{F(x) - F(c)}{x - c} \to g'(f(c))f'(c) \text{ as } x \to c.$$

Thus F is differentiable at c, with derivative

$$F'(c) = g'(f(c))f'(c).$$

■

If you are interested in why we need to allow for the possibility that $y = d$ in the above proof, here is an example of a situation where this arises. Suppose that f is the function

$$f(x) = \begin{cases} x^2\cos(1/x), & x \neq 0, \\ 0, & x = 0, \end{cases}$$

which is differentiable at 0, as you showed in Exercise F15(b). Then if $c = 0$ we have $d = f(c) = 0$, and since $\cos(n + \frac{1}{2})\pi = 0$ for $n \in \mathbb{N}$, it follows that $y = f(x)$ takes the value 0 infinitely many times in any interval containing 0.

Worked Exercise F13

Prove that each of the following composite functions is differentiable on its domain, and find its derivative.

(a) $k(x) = \sin(\cos x) \quad (x \in \mathbb{R})$

(b) $k(x) = \cosh(e^{2x}) \quad (x \in \mathbb{R})$

(c) $k(x) = \tan(x^2) \quad (x \in (-1, 1))$

Solution

(a) Here $k(x) = \sin(\cos x)$, so let

$$f(x) = \cos x \quad \text{and} \quad g(x) = \sin x \quad (x \in \mathbb{R}).$$

Then f and g are differentiable on $\mathbb{R}$, and

$$f'(x) = -\sin x \quad \text{and} \quad g'(x) = \cos x \quad (x \in \mathbb{R}).$$

Here we write out all the steps in the application of the Composition Rule, but it is not necessary for you to do this in your solutions.

By the Composition Rule, $k = g \circ f$ is differentiable on $\mathbb{R}$, and

$$\begin{aligned} k'(x) &= g'(f(x))f'(x) \\ &= \cos(\cos x) \times (-\sin x) \\ &= -\cos(\cos x)\sin x. \end{aligned}$$

(b) Here $k(x) = \cosh(e^{2x})$, so let

$$f(x) = 2x, \quad g(x) = e^x \quad \text{and} \quad h(x) = \cosh x \quad (x \in \mathbb{R}).$$

Then f, g and h are differentiable on $\mathbb{R}$, and

$$f'(x) = 2, \quad g'(x) = e^x \quad \text{and} \quad h'(x) = \sinh x \quad (x \in \mathbb{R}).$$

Here we use the extended form of the Composition Rule.

By the Composition Rule, $k = h \circ g \circ f$ is differentiable on $\mathbb{R}$, and

$$\begin{aligned} k'(x) &= h'(g(f(x)))g'(f(x))f'(x) \\ &= \sinh(e^{2x}) \times e^{2x} \times 2 \\ &= 2e^{2x}\sinh(e^{2x}). \end{aligned}$$

(c) In the notation of the Composition Rule, we can write

$$f(x) = x^2 \quad (x \in I) \quad \text{and} \quad g(x) = \tan x,$$

where $I = (-1, 1)$. Then $f(I) = [0, 1)$, so if we choose $J = (-\pi/2, \pi/2)$, then $f(I) \subseteq J$, as required.

Now

$$f'(x) = 2x \quad (x \in (-1, 1))$$

and

$$g'(x) = \sec^2 x.$$

Thus, by the Composition Rule, $k = g \circ f$ is differentiable on $(-1, 1)$, and

$$\begin{aligned} k'(x) &= g'(f(x))f'(x) \\ &= 2x \sec^2(x^2) \quad (x \in (-1, 1)). \end{aligned}$$

Exercise F21

Find the derivative of each of the following functions.

(a) $f(x) = \sinh(x^2) \quad (x \in \mathbb{R})$

(b) $f(x) = \sin(\sinh 2x) \quad (x \in \mathbb{R})$

(c) $f(x) = \sin\left(\dfrac{\cos 2x}{x^2}\right) \quad (x \in (0, \infty))$

2.3 Inverse Function Rule

In Section 4 of Unit D4 we discussed inverse functions and proved that if a function f with domain an interval I and image set $J = f(I)$ is strictly monotonic and continuous on I, then J is an interval, and f possesses a strictly monotonic and continuous inverse function f^{-1} with domain J. (Recall that *strictly monotonic* means that f is either strictly increasing or strictly decreasing.) In particular, we showed that the power functions, the trigonometric functions, the exponential function and the hyperbolic functions all have continuous inverse functions, provided that we restrict their domains where necessary.

These standard functions are all differentiable on their domains, and we now investigate whether their inverse functions also have this property.

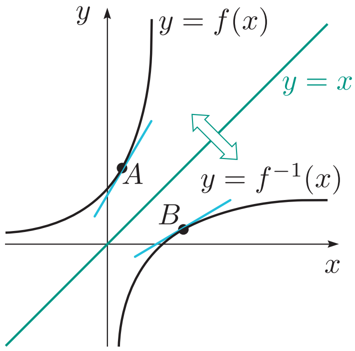


Figure 9 The tangent at a point on the graph $y = f(x)$ and its reflection in the line $y = x$

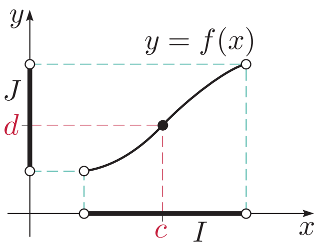


Figure 10 The graph of f in Theorem F33

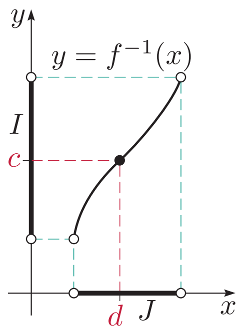


Figure 11 The graph of f^{-1} in Theorem F33

It is instructive to begin by considering their graphs. Recall that we obtain the graph $y = f^{-1}(x)$ by reflecting the graph $y = f(x)$ in the line $y = x$, which maps a typical point $A(c, d)$ on the graph $y = f(x)$ to the point $B(d, c)$ on the graph $y = f^{-1}(x)$, as illustrated in Figure 9. This suggests that if the gradient of the tangent to the graph $y = f(x)$ at the point A is $f'(c) = m$, then the gradient of the tangent to the graph $y = f^{-1}(x)$ at B is $(f^{-1})'(d) = 1/m$, provided that $m \neq 0$. However, if the graph of f has a horizontal tangent $(m = 0)$ at a point A, then the graph of f^{-1} has a vertical tangent at the corresponding point B; in this case, f^{-1} is not differentiable at B, since $1/m$ is not defined for $m = 0$. We therefore need the condition '$f'(x)$ is non-zero' in our statement of the rule for differentiating inverse functions. The Inverse Function Rule for differentiable functions is formally stated as the next theorem and illustrated in Figures 10 and 11.

Theorem F33 Inverse Function Rule for differentiable functions

Let f be a function whose domain is an open interval I on which f is continuous and strictly monotonic. Then f has an inverse function f^{-1} with domain $J = f(I)$.

If f is differentiable on I and $f'(x) \neq 0$ for $x \in I$, then f^{-1} is differentiable on J. Also, if $c \in I$ and $d = f(c)$, then

$$(f^{-1})'(d) = \frac{1}{f'(c)}.$$

The Leibniz notation for derivatives can be used to express the Inverse Function Rule in a form that is easy to remember: if we put

$$y = f(x) \quad \text{and} \quad x = f^{-1}(y),$$

and write

$$\frac{dy}{dx} \text{ for } f'(x) \quad \text{and} \quad \frac{dx}{dy} \text{ for } (f^{-1})'(y),$$

then

$$\frac{dx}{dy} = \frac{1}{dy/dx}.$$

Proof of Theorem F33 First note that f has an inverse function f^{-1} with domain $J = f(I)$ by the Inverse Function Theorem for continuous functions, proved in Subsection 4.1 of Unit D4.

Let $y \in J$ and $y \neq d$, so $f^{-1}(y) = x$, where $x \in I$ and $x \neq c$ (since f is strictly monotonic).

Then the difference quotient for f^{-1} at d is

$$\begin{aligned}\frac{f^{-1}(y)-f^{-1}(d)}{y-d} &= \frac{x-c}{f(x)-f(c)}\\ &= 1\Big/\frac{f(x)-f(c)}{x-c}.\end{aligned}$$

As $y \to d$, we have $x = f^{-1}(y) \to c$, since f^{-1} is continuous. So

$$\begin{aligned}\frac{f^{-1}(y)-f^{-1}(d)}{y-d} &= 1\Big/\frac{f(x)-f(c)}{x-c}\\ &\to 1/f'(c) \text{ as } y \to d \quad (\text{since } f'(c) \neq 0).\end{aligned}$$

Thus f^{-1} is differentiable at d, with derivative $(f^{-1})'(d) = 1/f'(c)$. So f^{-1} is differentiable on J. ■

The next worked exercise shows how the Inverse Function Rule can be used to determine the derivative of the inverse for some standard functions. (The derivatives of these inverse functions are included in the table of standard derivatives given at the end of this unit and in the module Handbook.)

Worked Exercise F14

For each of the following functions f, state the domain and rule of f^{-1}, show that f^{-1} is differentiable and determine its derivative.

(a) $f(x) = x^n \quad (x \in \mathbb{R}^+)$, where $n \in \mathbb{N}$, $n \geq 2$

(b) $f(x) = \tan x \quad (x \in (-\pi/2, \pi/2))$

(c) $f(x) = e^x \quad (x \in \mathbb{R})$

Solution

(a) The function

$$f(x) = x^n \quad (x \in \mathbb{R}^+)$$

is continuous and strictly increasing, and $f((0,\infty)) = (0,\infty)$. Also, f is differentiable on $(0,\infty)$, and its derivative $f'(x) = nx^{n-1}$ is non-zero there. So f satisfies the conditions of the Inverse Function Rule.

Hence f has an inverse function f^{-1} and $f^{-1}(y) = y^{1/n}$ is differentiable on its domain $(0,\infty)$. If $y = f(x) = x^n$ (so that $x = y^{1/n}$), then

$$(f^{-1})'(y) = \frac{1}{f'(x)} = \frac{1}{nx^{n-1}} = \frac{1}{n(y^{1/n})^{n-1}} = \frac{1}{n}\,y^{(1/n)-1}.$$

Replacing the domain variable y by x, we obtain

$$(f^{-1})'(x) = \frac{1}{n}x^{(1/n)-1} \quad (x \in (0,\infty)).$$

(b) The function

$$f(x) = \tan x \quad (x \in (-\pi/2, \pi/2))$$

is continuous and strictly increasing, and $f((-\pi/2, \pi/2)) = \mathbb{R}$. Also, f is differentiable on $(-\pi/2, \pi/2)$, and its derivative $f'(x) = \sec^2 x$ is non-zero there. So f satisfies the conditions of the Inverse Function Rule.

Hence f has an inverse function f^{-1} and $f^{-1}(y) = \tan^{-1} y$ is differentiable on its domain $\mathbb{R}$. If $y = f(x) = \tan x$, then

$$(f^{-1})'(y) = \frac{1}{f'(x)} = \frac{1}{\sec^2 x} = \frac{1}{1 + \tan^2 x} = \frac{1}{1 + y^2}.$$

Replacing the domain variable y by x, we obtain

$$(\tan^{-1})'(x) = \frac{1}{1 + x^2} \quad (x \in \mathbb{R}).$$

(c) The function

$$f(x) = e^x \quad (x \in \mathbb{R})$$

is continuous and strictly increasing, and $f(\mathbb{R}) = (0, \infty)$. Also, f is differentiable on $\mathbb{R}$, and its derivative $f'(x) = e^x$ is non-zero there. So f satisfies the conditions of the Inverse Function Rule.

Hence f has an inverse function f^{-1} (which we call log) and $f^{-1}(y) = \log y$ is differentiable on its domain $(0, \infty)$. If $y = f(x) = e^x$, then

$$(f^{-1})'(y) = \frac{1}{f'(x)} = \frac{1}{e^x} = \frac{1}{y}.$$

Replacing the domain variable y by x, we obtain

$$(\log)'(x) = \frac{1}{x} \quad (x \in (0, \infty)).$$

Exercise F22

For each of the following functions f, show that f^{-1} is differentiable and determine its derivative.

(a) $f(x) = \cos x \quad (x \in (0, \pi))$ (b) $f(x) = \sinh x \quad (x \in \mathbb{R})$

The Inverse Function Rule can sometimes be used to find values of the derivative of an inverse function f^{-1} even when the equation $y = f(x)$ cannot be solved to give a formula for the rule of f^{-1}. The next exercise illustrates this point.

Exercise F23

Let $f(x) = x^5 + x - 1$ $(x \in \mathbb{R})$.

(a) Prove that f has an inverse function f^{-1} which is differentiable on $\mathbb{R}$. (You may assume here that $f(\mathbb{R}) = \mathbb{R}$; this was proved in Worked Exercise D54 in Unit D4.)

(b) Find the values of $(f^{-1})'(d)$ at those points $d = f(c)$ where $c = 0, 1, -1$.

Exponential functions

In Subsection 4.3 of Unit D4 we defined the number a^x, for $a > 0$, by the formula

$$a^x = \exp(x \log a).$$

Since the functions exp and log are differentiable on $\mathbb{R}$ and $\mathbb{R}^+$, respectively (by Theorem F21(e) and Worked Exercise F14(c)), it follows that we can use this formula to determine the derivatives of several related functions. The functions in the next two worked exercises are included in the table of standard derivatives. Notice that the derivative in Worked Exercise F15 agrees with the formula for the derivative of $f(x) = x^n$, where $n \in \mathbb{N}$.

Worked Exercise F15

Prove that, for $\alpha \in \mathbb{R}$, the function

$$f(x) = x^\alpha \quad (x \in \mathbb{R}^+)$$

is differentiable on its domain, and that

$$f'(x) = \alpha x^{\alpha - 1} \quad (x \in \mathbb{R}^+).$$

Solution

By definition,

$$f(x) = \exp(\alpha \log x) \quad (x \in \mathbb{R}^+).$$

The function $x \longmapsto \alpha \log x$ is differentiable on $\mathbb{R}^+$, with derivative α/x. Thus, by the Composition Rule, f is differentiable on $\mathbb{R}^+$, with derivative

$$\begin{aligned} f'(x) &= \exp(\alpha \log x) \times (\alpha/x) \\ &= x^\alpha(\alpha/x) = \alpha x^{\alpha-1} \quad (x \in \mathbb{R}^+). \end{aligned}$$

Worked Exercise F16

Prove that, for $a > 0$, the function

$$f(x) = a^x \quad (x \in \mathbb{R})$$

is differentiable, and that

$$f'(x) = a^x \log a \quad (x \in \mathbb{R}).$$

Solution

By definition,

$$f(x) = \exp(x \log a) \quad (x \in \mathbb{R}).$$

The function $x \longmapsto x \log a$ is differentiable on $\mathbb{R}$, with derivative $\log a$. By the Composition Rule, f is differentiable on $\mathbb{R}$, with derivative

$$\begin{aligned} f'(x) &= \exp(x \log a) \times \log a \\ &= a^x \log a. \end{aligned}$$

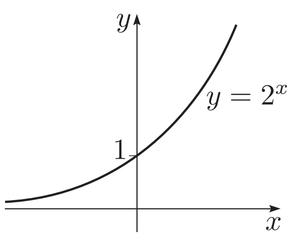


Figure 12 The graph of $f(x) = 2^x$

At the beginning of Book D we posed the following question: does the graph of the function $f(x) = 2^x$, shown in Figure 12, have a jump at the point $\sqrt{2}$? We showed in Subsection 4.3 of Unit D4 that f is continuous, which proves that its graph has no jumps. In Worked Exercise F16 we have now shown that f also is differentiable, so its graph has a tangent at every point and thus has no corners.

Exercise F24

Prove that the function

$$f(x) = x^x \quad (x \in \mathbb{R}^+)$$

is differentiable, and find its derivative.

3 Rolle's Theorem

In this section and in Section 4 you will meet some of the fundamental properties of functions that are differentiable not just at a particular point, but *on an interval*. These results are motivated by the geometric significance of differentiability in terms of tangents, and they explain why the graphs of differentiable functions possess certain geometric properties.

3.1 Local Extreme Value Theorem

In Section 3 of Unit D4 you met some of the fundamental properties of functions which are continuous on a bounded closed interval. In particular, you studied the Extreme Value Theorem, which states that if a function f is continuous on a closed interval $[a, b]$, then there are points c and d in $[a, b]$ such that

$$f(x) \leq f(d), \quad \text{for } x \in [a, b],$$

and

$$f(x) \geq f(c), \quad \text{for } x \in [a, b].$$

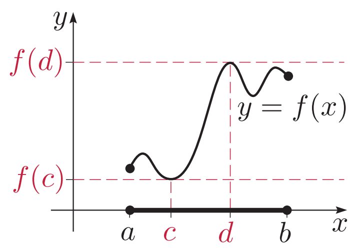


Figure 13 The extreme values of a function continuous on a closed interval

This is illustrated in Figure 13. The value $f(d)$ is the **maximum** of f on $[a, b]$, and the value $f(c)$ is the **minimum** of f on $[a, b]$. A maximum or a minimum of f is called an **extreme value** of f. But how do we determine the points c and d where these extreme values occur? In general, this is not easy. However, if the function f is differentiable, then we can, in principle, determine c and d by first finding any *local* extreme values of the function f on the interval $[a, b]$.

Roughly speaking, for a point c in (a, b), the value $f(c)$ is a *local maximum* of f on $[a, b]$ if $f(c)$ is the greatest value of f in the immediate vicinity of c, and a *local minimum* of f on $[a, b]$ if $f(c)$ is the least value of f in the immediate vicinity of c. These ideas are illustrated in Figure 14 and stated formally in the definitions which follow.

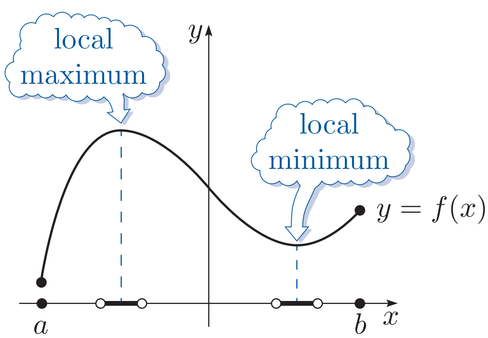


Figure 14 The local extreme values of a function

Definitions

The function f has

- a **local maximum** $f(c)$ at c if there is an open interval $I = (c - r, c + r)$, where $r > 0$, in the domain of f such that
$$f(x) \le f(c), \quad \text{for } x \in I$$
- a **local minimum** $f(c)$ at c if there is an open interval $I = (c - r, c + r)$, where $r > 0$, in the domain of f such that
$$f(x) \ge f(c), \quad \text{for } x \in I$$
- a **local extreme value** $f(c)$ at c if $f(c)$ is either a local maximum or a local minimum.

By definition, a local extreme value of a function f defined on a bounded closed interval $[a, b]$ is an interior point of $[a, b]$; that is, a local extreme value cannot occur at either of the endpoints a and b.

If we want to find the local extreme values of a differentiable function f, then we can use the following result, which gives a connection between the local extreme values of a function f and the points c where $f'(c) = 0$. A point c such that $f'(c) = 0$ is called a **stationary point** of f; we sometimes say that f' **vanishes** at c.

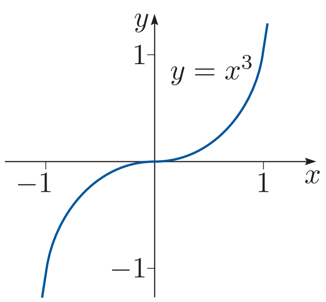


Figure 15 The graph of $f(x) = x^3$

Theorem F34 Local Extreme Value Theorem

If f has a local extreme value at c and f is differentiable at c, then
$$f'(c) = 0.$$

Note that the converse of the Local Extreme Value Theorem is *false*: a point where the derivative vanishes is not necessarily a local extreme value. For example, the function $f(x) = x^3$ does not have a local extreme value at 0, although $f'(0) = 0$; see Figure 15.

Proof of Theorem F34 We prove the result only for a local maximum; the proof of the local minimum version is similar. Suppose that f has a local maximum at c. Then there exists a positive number r such that
$$f(x) \le f(c), \quad \text{for } c - r < x < c + r. \tag{7}$$

Now let
$$x_n = c + \frac{r}{n} \quad \text{and} \quad x'_n = c - \frac{r}{n}, \quad n = 2, 3, \ldots.$$

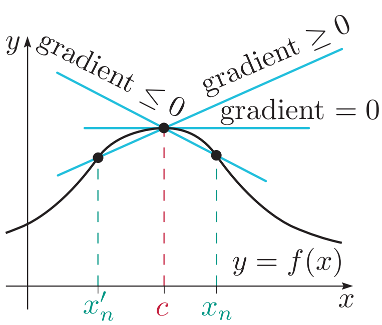


Figure 16 The proof of the Local Extreme Value Theorem

We do not include $n = 1$ as we require x_n and x'_n to lie in the open interval $(c - r, c + r)$. The argument which follows is illustrated in Figure 16.

Then $c < x_n < c + r$, for $n = 2, 3, \ldots$, so

$$f(x_n) - f(c) \le 0 \quad \text{and} \quad x_n - c > 0, \quad \text{for } n = 2, 3, \ldots,$$

by inequality (7). Hence

$$\frac{f(x_n) - f(c)}{x_n - c} \le 0, \quad \text{for } n = 2, 3, \ldots.$$

Since $x_n \to c$, we deduce by the Limit Inequality Rule (see Subsection 3.3 of Unit D2 *Sequences*) that

$$f'(c) = \lim_{n \to \infty} \frac{f(x_n) - f(c)}{x_n - c} \le 0. \tag{8}$$

On the other hand, $c - r < x'_n < c$, for $n = 2, 3, \ldots$, so

$$f(x'_n) - f(c) \le 0 \quad \text{and} \quad x'_n - c < 0, \quad \text{for } n = 2, 3, \ldots,$$

by inequality (7). Hence

$$\frac{f(x'_n) - f(c)}{x'_n - c} \ge 0, \quad \text{for } n = 2, 3, \ldots.$$

Since $x'_n \to c$, we deduce that

$$f'(c) = \lim_{n \to \infty} \frac{f(x'_n) - f(c)}{x'_n - c} \ge 0. \tag{9}$$

Hence, by inequalities (8) and (9), we have $f'(c) = 0$, as required. ■

Any extreme value of a function f on a bounded closed interval $[a, b]$ which is not $f(a)$ or $f(b)$ must also be a local extreme value. Thus, by Theorem F34, such a point x must satisfy $f'(x) = 0$. This gives the following property of the extreme values of a differentiable function on a bounded closed interval.

Corollary F35

Let f be continuous on the closed interval $[a, b]$ and differentiable on (a, b). Then the extreme values of f on $[a, b]$ can occur only at a or b, or at points x in (a, b) where $f'(x) = 0$.

We now reformulate Corollary F35 as a strategy for locating maxima and minima.

Strategy F6

To find the maximum and minimum of a function f that is continuous on $[a, b]$ and differentiable on (a, b), do the following.

1. Determine the points $c_1, c_2, \ldots$ in (a, b) where f' is zero.
2. Hence determine the values of

   $$f(a),\ f(b),\ f(c_1),\ f(c_2),\ \ldots;$$

   the greatest of these is the maximum and the least is the minimum.

Note that, in some cases, there may be infinitely many points in (a, b) where f' is zero; this is so, for example, if f is constant on (a, b).

Exercise F25

Use Strategy F6 to determine the maximum and minimum of the function

$$f(x) = \sin^2 x + \cos x$$

on the interval $[0, \pi/2]$.

3.2 Rolle's Theorem

In the last subsection you saw that if a function f is continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b), then the extreme values of f can occur only at a or b, or at some point c in (a, b) such that $f'(c) = 0$. If we also know that the values of $f(a)$ and $f(b)$ are equal, then we can deduce that there must be some point c in (a, b) where $f'(c) = 0$. This is illustrated in Figure 17.

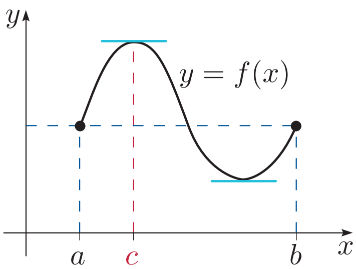


Figure 17 A function with $f(a) = f(b)$

Theorem F36 Rolle's Theorem

Let f be continuous on the closed interval $[a, b]$ and differentiable on (a, b). If $f(a) = f(b)$, then there exists a point c, with $a < c < b$, such that

$$f'(c) = 0.$$

Remarks

1. Rolle's Theorem is one of the most important theorems in analysis; for example, most of the results you will meet in Sections 4 and 5 depend on Rolle's Theorem.
2. Rolle's Theorem is an *existence theorem*; that is, it tells us that a point c with the stated property exists, but not how to find it. Often it is difficult to evaluate c explicitly, and there may be more than one point c in (a, b) at which f' vanishes.
3. In geometric terms, Rolle's Theorem states that if the line joining the points $(a, f(a))$, $(b, f(b))$ on the graph of f is horizontal, then so is the tangent to the graph for some $c \in (a, b)$.

Rolle's Theorem is one of the foundational theorems in differential calculus. Its importance lies in the fact that it is needed in the proof of the Mean Value Theorem and for establishing the existence of Taylor series. However, when Michel Rolle (1652–1719) made the first statement of the theorem in 1690, Taylor series had not yet been discovered and calculus was still in its infancy. Moreover, Rolle was deeply suspicious of its methods. The first appearance of his theorem was not in the context of calculus at all, but in the context of solving equations. It is paradoxical that the name of a man renowned for his opposition to the infinitesimal calculus should end up attached to one of the fundamental theorems in the subject.

During the eighteenth century the theorem lived on as a theorem in algebra until it appeared in the work of Leonhard Euler (1707–1783) in 1755. Euler, with the calculus at his fingertips, had no need for Rolle's rather convoluted algebraic method and so for the first time the theorem resembled its modern counterpart. It is not known whether Euler had read Rolle's work since, characteristically, he made no reference to any earlier work. By the 1830s the theorem had appeared in a number of textbooks on the theory of equations, but was still not associated with Rolle. It was first ascribed to Rolle by Wilhelm Drobisch (1802–1896), a professor at the University of Leipzig, in a textbook of 1834.

In the latter half of the nineteenth century the theorem underwent its second significant change. From being a useful result in the theory of equations it was transformed into a fundamental theorem in analysis. In 1873 Charles Hermite (1822–1901) used the theorem in his *Cours d'Analyse* in the context of the theory of Taylor series, clearly attributing it to Rolle. Hermite was the leading French analyst of his generation and his *Cours d'Analyse* was extremely influential in France during the latter part of the nineteenth century. His unequivocal association of the theorem with Rolle was decisive for future writers.

Charles Hermite

Proof of Theorem F36 Suppose that $f(a) = f(b)$.

If f is constant on $[a, b]$, then $f'(x) = 0$ everywhere on (a, b); in this case, we can take c to be any point of (a, b).

If f is non-constant on $[a, b]$, then either the maximum or the minimum (or both) of f on $[a, b]$ is different from the common value $f(a) = f(b)$.

Since f is continuous on $[a, b]$, it must have both a maximum and a minimum on $[a, b]$, by the Extreme Value Theorem; see Subsection 3.3 of Unit D4.

Since one of the extreme values occurs at some point c with $a < c < b$, the Local Extreme Value Theorem shows that $f'(c)$ must be zero. ■

We can use Rolle's Theorem to verify the existence of zeros of certain functions which are derivatives.

Worked Exercise F17

Use Rolle's Theorem to show that if

$$f(x) = 3x^4 - 2x^3 - 2x^2 + 2x,$$

then there is a value of c in $(-1, 1)$ such that $f'(c) = 0$.

Solution

Since f is a polynomial function, f is continuous on $[-1, 1]$ and differentiable on $(-1, 1)$. Also, $f(1) = f(-1) = 1$. Thus f satisfies the conditions of Rolle's Theorem on $[-1, 1]$. It follows that there exists a point $c \in (-1, 1)$ such that $f'(c) = 0$.

For the function in Worked Exercise F17 we can find a value for c directly by using the fact that

$$\begin{aligned} f'(x) &= 12x^3 - 6x^2 - 4x + 2 \\ &= 2(3x^2 - 1)(2x - 1). \end{aligned}$$

Thus f' has a zero at each of the points $-1/\sqrt{3}$, $1/\sqrt{3}$ and $\frac{1}{2}$, which are all in $(-1, 1)$.

Exercise F26

Use Rolle's Theorem to show that if

$$f(x) = x^4 - 4x^3 + 3x^2 + 2,$$

then there is a value of c in $(1, 3)$ such that $f'(c) = 0$.

Exercise F27

For each of the following functions, state whether Rolle's Theorem applies for the given interval.

(a) $f(x) = \tan x, \quad [0, \pi]$

(b) $f(x) = 3|x - 1| - x, \quad [0, 3]$

(c) $f(x) = x - 9x^{17} + 8x^{18}, \quad [0, 1]$

(d) $f(x) = \sin x + \tan^{-1} x, \quad [0, \pi/2]$

4 Mean Value Theorem

In this section you will continue to study the geometric properties of functions that are differentiable on intervals and meet some of their applications.

4.1 Mean Value Theorem

First we recall the geometric interpretation of Rolle's Theorem from the previous section. Rolle's Theorem tells us that, under suitable conditions, if the chord joining the points $(a, f(a))$ and $(b, f(b))$ on the graph of f is horizontal, then so is the tangent to the graph for some c in (a, b).

If you imagine pushing this horizontal chord, always parallel to its original position, until it is just about to lose contact with the graph of f, then it appears that at this point the chord is a tangent to the graph; see the top graph in Figure 18. This 'chord-pushing' approach suggests that even if the original chord is not horizontal (that is, if $f(a) \neq f(b)$), then there must still be a point c in (a, b) at which the tangent is parallel to the chord; see the bottom graph in Figure 18.

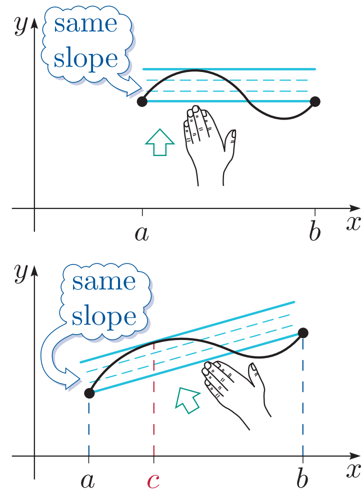


Figure 18 A tangent parallel to a chord

Worked Exercise F18

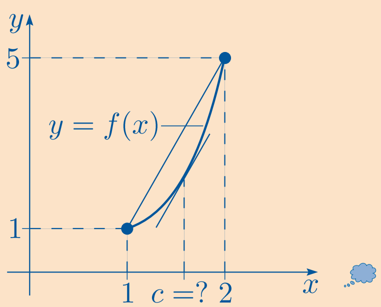
Consider the function

$$f(x) = x^3 - 3x + 3 \quad (x \in [1, 2]).$$

Find a point $c \in (1, 2)$ such that the tangent to the graph of f is parallel to the chord joining $(1, f(1))$ to $(2, f(2))$.

Solution

The graph of f is shown below.



Since $f(1) = 1 - 3 + 3 = 1$ and $f(2) = 8 - 6 + 3 = 5$, the gradient of the chord is

$$\frac{f(2) - f(1)}{2 - 1} = \frac{5 - 1}{2 - 1} = 4.$$

Now f is a polynomial, so it is differentiable on $(1, 2)$, and its derivative is $f'(x) = 3x^2 - 3$. Hence $f'(c) = 4$ when $3c^2 = 7$; that is, when $c = \sqrt{7/3} \simeq 1.53$.

Thus at the point $(c, f(c))$ the tangent to the graph is parallel to the chord joining the endpoints of the graph.

We now generalise Rolle's Theorem and show that there is always a point where the tangent to the graph is parallel to the chord joining the endpoints. This result is known as the *Mean Value Theorem*, so-called since the gradient of the chord,

$$\frac{f(b) - f(a)}{b - a},$$

can be thought of as the *mean value* of the derivative between a and b.

Theorem F37 Mean Value Theorem

Let f be continuous on the closed interval $[a, b]$ and differentiable on (a, b). Then there exists a point c in (a, b) such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Remarks

1. Like Rolle's Theorem, this result is an *existence theorem*: it does not tell us what the point c is, just that such a point exists.
2. When $f(a) = f(b)$, the Mean Value Theorem reduces to Rolle's Theorem.

Proof of Theorem F37 The gradient of the chord joining the points $(a, f(a))$ and $(b, f(b))$ is

$$m = \frac{f(b) - f(a)}{b - a},$$

so the equation of the chord is

$$y = m(x - a) + f(a).$$

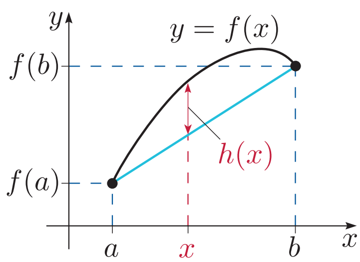


Figure 19 The vertical distance $h(x)$

It follows that, for $x \in [a, b]$, the vertical distance $h(x)$ from the point $(x, f(x))$ to the chord, as shown in Figure 19, is given by the function

$$h(x) = f(x) - (m(x - a) + f(a)).$$

Now $h(a) = h(b) = 0$, and h is continuous on $[a, b]$ and differentiable on (a, b). Thus h satisfies all the conditions of Rolle's Theorem.

It follows that there exists a point c in (a,b) such that $h'(c)=0$. Since $h'(c)=f'(c)-m$, we have

$$f'(c)=m=\frac{f(b)-f(a)}{b-a},$$

as required. ■

Worked Exercise F19

Use the Mean Value Theorem to show that if

$$f(x)=\frac{x-1}{x+1},$$

then there is a point c in $\left(1,\frac{7}{2}\right)$ such that $f'(c)=\frac{2}{9}$.

Solution

The function f is a rational function whose denominator is non-zero on $\left[1,\frac{7}{2}\right]$, so f is continuous on $\left[1,\frac{7}{2}\right]$ and differentiable on $\left(1,\frac{7}{2}\right)$. Thus f satisfies the conditions of the Mean Value Theorem.

Now

$$\frac{f\left(\frac{7}{2}\right)-f(1)}{\frac{7}{2}-1}=\frac{\frac{5}{9}-0}{\frac{5}{2}}=\frac{2}{9}.$$

Thus, by the Mean Value Theorem, there exists a point c in $\left(1,\frac{7}{2}\right)$ such that $f'(c)=\frac{2}{9}$.

For the function in Worked Exercise F19 we can find a value for c directly. Since $f'(x)=2/(x+1)^2$, the point c satisfies

$$f'(c)=\frac{2}{(c+1)^2}=\frac{2}{9};\quad\text{that is,}\quad (c+1)^2=9.$$

This equation has solutions 2 and -4, so $c=2$ $\left(\text{since } 2\in\left(1,\frac{7}{2}\right)\right)$.

In the following exercise, a value for c cannot be found in this direct way.

Exercise F28

Use the Mean Value Theorem to show that if

$$f(x)=xe^x,$$

then there is a point c in $(0,2)$ such that $f'(c)=e^2$.

4.2 Positive, negative and zero derivatives

We now study some consequences of the Mean Value Theorem for functions whose derivatives are always positive, always negative, or always zero. First we prove a fundamental result about monotonic functions, which you used to help with graph sketching in Section 2 of Unit A4 *Real functions, graphs and conics.*

For any interval I, the **interior** of I is the largest open subinterval of I. It is obtained from I by removing any endpoints of I, so it consists of all the interior points of I.

Theorem F38 Increasing–Decreasing Theorem

Let f be continuous on an interval I and differentiable on the interior of I.

(a) If $f'(x) \geq 0$ for x in the interior of I, then f is increasing on I.

(b) If $f'(x) \leq 0$ for x in the interior of I, then f is decreasing on I.

Proof Choose any two points x_1 and x_2 in I, with $x_1 < x_2$. The function f satisfies the conditions of the Mean Value Theorem on the interval $[x_1, x_2]$, so there exists a point c in (x_1, x_2) such that

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(c).$$

Hence $f(x_2) - f(x_1)$ has the same sign as $f'(c)$, so

(a) if $f'(x) \geq 0$ for x in the interior of I, then $f(x_2) - f(x_1) \geq 0$ and hence f is increasing on I

(b) if $f'(x) \leq 0$ for x in the interior of I, then $f(x_2) - f(x_1) \leq 0$ and hence f is decreasing on I. ■

There is also a version of Theorem F38 with the weak inequalities replaced by strict inequalities, in which case the conclusions are as follows.

(a) If $f'(x) > 0$ for x in the interior of I, then f is strictly increasing on I.

(b) If $f'(x) < 0$ for x in the interior of I, then f is strictly decreasing on I.

Note that the converse of each of these two statements with strict inequalities is *false*. For example, $f(x) = x^3$ is strictly increasing on $[-1, 1]$, but $f'(0) = 0$.

Exercise F29

On the given interval I, determine whether each of the following functions f is

- strictly increasing
- increasing, but not strictly increasing
- strictly decreasing, or
- decreasing, but not strictly decreasing.

(a) $f(x) = 3x^{4/3} - 4x, \quad I = [1, \infty)$

(b) $f(x) = x - \log x, \quad I = (0, 1]$

The following corollary to Theorem F38 will be useful in later units.

Corollary F39 Zero Derivative Theorem

Let f be continuous on an interval I and differentiable on the interior of I. If $f'(x) = 0$ for all x in the interior of I, then f is constant on I.

Proof Theorem F38(a) and (b) both apply, so f is both increasing and decreasing on I. Hence f is constant on I. ■

We can often determine whether a point c such that $f'(c) = 0$ is a local maximum or a local minimum of a function f by using the following test.

Theorem F40 Second Derivative Test

Let f be a twice-differentiable function defined on an open interval I containing a point c such that $f'(c) = 0$ and f'' is continuous at c.

(a) If $f''(c) > 0$, then $f(c)$ is a local minimum of f.

(b) If $f''(c) < 0$, then $f(c)$ is a local maximum of f.

Proof We prove part (a); the proof of part (b) is similar, so we omit it.

We first use the ε-δ definition of continuity with $\varepsilon = \frac{1}{2}f''(c)$ to show that $f''(x) > 0$ for x close to c. A similar technique was used in the proof of the Quotient Rule in Subsection 2.1.

Suppose that $f''(c) > 0$. Since f'' is continuous at c, there exists $\delta > 0$ such that $(c - \delta, c + \delta) \subseteq I$ and

$$|f''(x) - f''(c)| < \tfrac{1}{2}f''(c), \quad \text{for } x \in (c - \delta, c + \delta),$$

so

$$f''(x) > \tfrac{1}{2}f''(c) > 0, \quad \text{for } x \in (c - \delta, c + \delta).$$

Thus f' is strictly increasing on the open interval $(c-\delta, c+\delta)$, by the strict inequalities version of the Increasing–Decreasing Theorem. Since $f'(c) = 0$, we deduce that

$$\begin{aligned} f'(x) &< 0, \quad \text{for } x \in (c-\delta, c), \\ f'(x) &> 0, \quad \text{for } x \in (c, c+\delta). \end{aligned}$$

Thus f has a local minimum at c, by another application of the strict inequalities version of the Increasing–Decreasing Theorem. ■

Note that if $f''(c) = 0$, then the Second Derivative Test gives us no information about local extreme values. For example, the function $f(x) = x^3$ satisfies $f'(0) = 0$ and $f''(0) = 0$, but it has neither a local maximum nor a local minimum at 0.

Exercise F30

Consider the function

$$f(x) = x^3 - 3x^2 + 1.$$

(a) Determine those points c such that $f'(c) = 0$.

(b) Using the Second Derivative Test, determine whether the points c found in part (a) correspond to local maxima or local minima, and find the values of these local maxima or local minima.

Proving inequalities

We now demonstrate how the Increasing–Decreasing Theorem can be used to prove certain inequalities involving differentiable functions.

First we prove a generalisation of Bernoulli's Inequality. Recall from Subsection 3.5 of Unit D1 that Bernoulli's Inequality states that

$$(1+x)^n \geq 1 + nx, \quad \text{for } x \geq -1 \text{ and } n \in \mathbb{N}.$$

In the next worked exercise we show that this inequality still holds if we replace n by any real number $\alpha \geq 1$.

Worked Exercise F20

Let $\alpha \geq 1$. Prove that

$$(1+x)^\alpha \geq 1 + \alpha x, \quad \text{for } x \geq -1.$$

Solution

The case $\alpha = 1$ holds by Bernoulli's Inequality, so we can assume that $\alpha > 1$. Define the function

$$f(x) = (1+x)^\alpha - (1+\alpha x) \quad (x \in [-1,\infty)).$$

We want to show that $f(x) \geq 0$ for $x \in [-1,\infty)$. Since $f(0) = 1 - 1 = 0$, this is equivalent to showing that

$$f(x) \geq f(0), \quad \text{for } x \in [-1,\infty).$$

We do this by showing that

$$f \text{ is increasing on } (0,\infty) \text{ and decreasing on } [-1,0). \tag{$*1$}$$

Now the function f is continuous on $[-1,\infty)$ and differentiable on $(-1,\infty)$, with derivative

$$\begin{aligned} f'(x) &= \alpha(1+x)^{\alpha-1} - \alpha \\ &= \alpha((1+x)^{\alpha-1} - 1), \quad \text{for } x \in (-1,\infty). \end{aligned} \tag{$*2$}$$

If $x > 0$, then $1 + x > 1$, so

$$(1+x)^{\alpha-1} > 1, \quad \text{for } x > 0.$$

Here we have used Rule 5 for rearranging inequalities from Section 2 of Unit D1:

$$\text{if } a, b \geq 0 \text{ and } p > 0, \text{ then } a > b \iff a^p > b^p.$$

We have used this with $p = \alpha - 1 > 0$.

Hence, by equation ($*2$),

$$f'(x) > 0, \quad \text{for } x > 0,$$

so f is increasing on $(0,\infty)$, by the Increasing–Decreasing Theorem.

Similarly, if $-1 < x < 0$, then $0 < 1 + x < 1$, so

$$(1+x)^{\alpha-1} < 1, \quad \text{for } -1 < x < 0.$$

Hence, by equation ($*2$),

$$f'(x) < 0, \quad \text{for } -1 < x < 0,$$

so f is decreasing on $(-1, 0)$. This proves statement ($*1$).

Worked Exercise F20 illustrates the following general strategy for using the Increasing–Decreasing Theorem to prove inequalities.

Strategy F7

To prove that $g(x) \geq h(x)$, for $x \in [a, b]$, carry out the following steps.

1. Let
$$f(x) = g(x) - h(x),$$
and show that f is continuous on $[a, b]$ and differentiable on (a, b).

2. Prove that

either $f(a) \geq 0$ and $f'(x) \geq 0$ for $x \in (a, b)$,

or $f(b) \geq 0$ and $f'(x) \leq 0$ for $x \in (a, b)$.

The diagrams in Figure 20 below illustrate why Strategy F7 works.

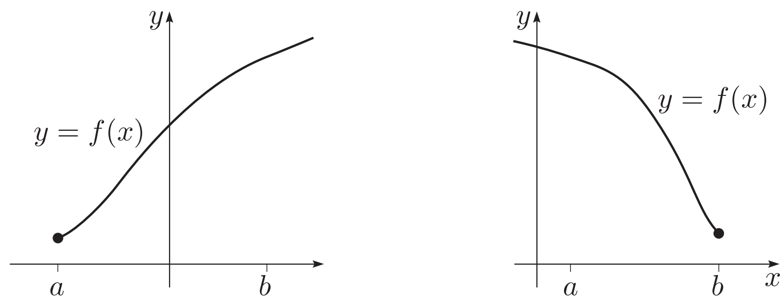


Figure 20 The two cases in Strategy F7

Remarks

1. There is a corresponding version of Strategy F7 in which the weak inequalities are replaced by strict inequalities.
2. We can also apply Strategy F7 to intervals of the form $[a, \infty)$ if the first case in step 2 holds, and to intervals of the form $(-\infty, b]$ if the second case in step 2 holds.
3. Notice that in Worked Exercise F20 we used *both* cases in step 2 of Strategy F7.

Exercise F31

Prove the following inequalities.

(a) $\sin x \leq x$, for $x \in [0, \infty)$ (b) $\frac{2}{3}x + \frac{1}{3} \geq x^{2/3}$, for $x \in [0, 1]$

5 L'Hôpital's Rule

In Section 1 we found the derivatives of sin and exp by using the results

$$\lim_{x\to 0}\frac{\sin x}{x}=1 \quad\text{and}\quad \lim_{x\to 0}\frac{e^x-1}{x}=1,$$

which you met in Section 1 of Unit F1. Each of the above limits is of the form

$$\lim_{x\to c}\frac{f(x)}{g(x)},$$

where f and g are continuous functions with $f(c)=g(c)=0$. Such limits cannot be evaluated by the Quotient Rule for limits, because this rule requires $\lim\limits_{x\to c} g(x)\neq 0$.

There are similar problems with the following more complicated limits:

$$\lim_{x\to \pi/2}\frac{\cos 3x}{\sin x-e^{\cos x}} \quad\text{and}\quad \lim_{x\to 0}\frac{x^2}{\cosh x-1}.$$

Do they exist? And if they do, what are their values?

In this section you will meet a result called l'Hôpital's Rule, which enables us to answer such questions.

5.1 Cauchy's Mean Value Theorem

In Section 4 you met the Mean Value Theorem which asserts that, under certain conditions, a function defined on a closed interval has the property that at some intermediate point, the tangent to its graph is parallel to the chord joining the endpoints. To prove l'Hôpital's Rule, we will need the following generalisation of the Mean Value Theorem.

Theorem F41 Cauchy's Mean Value Theorem

Let f and g be continuous on $[a,b]$ and differentiable on (a,b). Then there exists a point $c\in(a,b)$ such that

$$f'(c)(g(b)-g(a))=g'(c)(f(b)-f(a));$$

in particular, if $g(b)\neq g(a)$ and $g'(c)\neq 0$, then

$$\frac{f'(c)}{g'(c)}=\frac{f(b)-f(a)}{g(b)-g(a)}.$$

Remarks

1. Note that Cauchy's Mean Value Theorem involves *two* functions defined on a closed interval $[a, b]$ and, subject to the stated conditions, gives us an expression for the ratio of their derivatives at some point $c \in (a, b)$. (It is this expression that we need for the proof of l'Hôpital's Rule.) If we put $g(x) = x$, then Cauchy's Mean Value Theorem reduces to the usual Mean Value Theorem.
2. There is a geometric interpretation of Cauchy's Mean Value Theorem that you may find helpful. Recall from Subsection 5.4 of Unit A4 that we can describe a curve in $\mathbb{R}^2$ by specifying the x- and y-coordinates of its points using two functions, f and g, to define parametric equations

   $$x = g(t), \quad y = f(t),$$

   where the parameter t belongs to some suitable interval $[a, b]$. Thus we can think of the two functions in Cauchy's Mean Value Theorem as defining a curve in this way; see Figure 21.

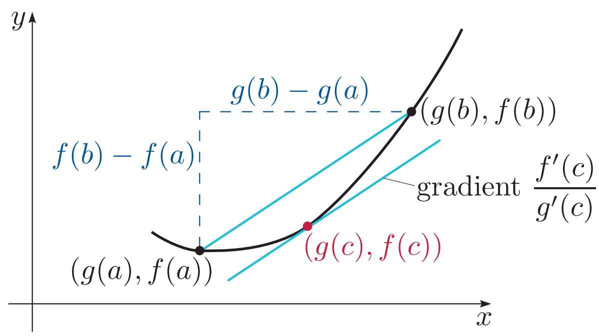
   

   Figure 21 A geometric interpretation of Cauchy's Mean Value Theorem

   Now it can be shown that, subject to certain conditions, the gradient of the tangent to this curve at the point $(g(c), f(c))$ is $\dfrac{f'(c)}{g'(c)}$ (though we do not prove this here). Thus, Cauchy's Mean Value Theorem tells us that there is some value c in the interval (a, b) such that the gradient of the curve at the point $(g(c), f(c))$ is equal to the gradient $\dfrac{f(b) - f(a)}{g(b) - g(a)}$ of the chord joining the endpoints of the curve, $(g(a), f(a))$ and $(g(b), f(b))$.

From 1815 to 1830 Augustin-Louis Cauchy (1789–1857) taught at the famous École Polytechnique in Paris, the École founded for the training of engineers. Cauchy wrote several textbooks on analysis designed for the students there, including his *Cours d'Analyse* of 1821 and his *Résumé des leçons données a l'École Royale Polytechnique sur le calcul infinitésimal* of 1823, the latter containing the result now known as the Cauchy (or Generalised) Mean Value Theorem. His aim, as he explained in the *Cours d'Analyse*, was to endow proof in analysis with the same level of rigour as proof in Euclid's geometry. However, the originality of his lectures within the programme of the École was not looked upon favourably, and he was criticised by the director of the École who warned Cauchy that 'It is the opinion of many persons that the instruction in pure mathematics is being carried too far at the École and such an uncalled-for extravagance is prejudicial to other branches of mathematics.'

Proof of Theorem F41 Consider the function

$$h(x) = f(x)(g(b) - g(a)) - g(x)(f(b) - f(a)).$$

By the Combination Rules for continuous functions and for differentiable functions, h is continuous on $[a, b]$ and differentiable on (a, b). Also,

$$\begin{aligned} h(a) &= f(a)(g(b) - g(a)) - g(a)(f(b) - f(a)) \\ &= f(a)g(b) - g(a)f(b) \end{aligned}$$

and

$$\begin{aligned} h(b) &= f(b)(g(b) - g(a)) - g(b)(f(b) - f(a)) \\ &= f(a)g(b) - g(a)f(b), \end{aligned}$$

so $h(a) = h(b)$.

Thus h satisfies the conditions of Rolle's Theorem on $[a, b]$, so there exists a point $c \in (a, b)$ for which

$$h'(c) = 0;$$

that is,

$$f'(c)(g(b) - g(a)) - g'(c)(f(b) - f(a)) = 0.$$

The two equations in the statement of Cauchy's Mean Value Theorem now follow by rearranging this equation. ■

5.2 L'Hôpital's Rule and its application

We are now in a position to prove the main result of this section.

Theorem F42 L'Hôpital's Rule

Let f and g be differentiable on an open interval I containing the point c, and suppose that $f(c) = g(c) = 0$. Then

$$\lim_{x\to c} \frac{f(x)}{g(x)} \text{ exists and equals } \lim_{x\to c} \frac{f'(x)}{g'(x)},$$

provided that the latter limit exists.

Proof We assume that

$$\lim_{x\to c} \frac{f'(x)}{g'(x)} \text{ exists and equals } l, \tag{10}$$

and we want to deduce that

$$\lim_{x\to c} \frac{f(x)}{g(x)} = l. \tag{11}$$

We use the ε-δ definition of limit from Subsection 3.3 of Unit F1.

Let $\varepsilon > 0$. Then, by statement (10), there exists $\delta > 0$ such that

$$\left| \frac{f'(x)}{g'(x)} - l \right| < \varepsilon, \quad \text{for all } x \text{ with } 0 < |x - c| < \delta. \tag{12}$$

In particular, $g'(x) \neq 0$, for $0 < |x - c| < \delta$.

Suppose now that x is such that $0 < |x - c| < \delta$. If $g(x) = g(c)$, then g' must vanish at some point between c and x (by Rolle's Theorem), which we know is not the case. Therefore, $g(x) \neq g(c)$. Thus, by Cauchy's Mean Value Theorem, there exists some point d between c and x such that

$$\begin{aligned}\frac{f'(d)}{g'(d)} &= \frac{f(x) - f(c)}{g(x) - g(c)} \\ &= \frac{f(x)}{g(x)} \quad (\text{since } f(c) = g(c) = 0).\end{aligned}$$

Thus, by statement (12), we have

$$\left| \frac{f(x)}{g(x)} - l \right| = \left| \frac{f'(d)}{g'(d)} - l \right| < \varepsilon, \quad \text{for all } x \text{ with } 0 < |x - c| < \delta.$$

It follows that statement (11) is true, as required. ■

In the early 1690s the Marquis de l'Hôpital (1661–1704) contracted Johann Bernoulli (1667–1748) to teach him the recently published Leibnizian differential calculus. The result was the first textbook ever written on the calculus, l'Hôpital's *Analyse des infiniment petits, pour l'intelligence des lignes courbes* (*The analysis of the infinitely small, for the understanding of curved lines*), published in Paris in 1696. It contains what it is now known as l'Hôpital's rule, although l'Hôpital learnt the rule from Bernoulli. When l'Hôpital published his *Analyse* he acknowledged all the instruction he had received from Bernoulli but in such a way that Bernoulli took offence.

The Marquis de l'Hôpital

We can use l'Hôpital's Rule to evaluate the two complicated limits mentioned in the introduction to this section.

Worked Exercise F21

Prove that

$$\lim_{x\to\pi/2} \frac{\cos 3x}{\sin x - e^{\cos x}}$$

exists, and determine its value.

Solution

Let $I = \mathbb{R}$ and define

$$f(x) = \cos 3x \quad \text{and} \quad g(x) = \sin x - e^{\cos x} \quad (x \in \mathbb{R}).$$

Then f and g are differentiable, and $f(\pi/2) = g(\pi/2) = 0$; hence f and g satisfy the conditions of l'Hôpital's Rule at the point $x = \pi/2$.

Now the derivatives of f and g are

$$f'(x) = -3\sin 3x \quad \text{and} \quad g'(x) = \cos x + e^{\cos x}\sin x \quad (x \in \mathbb{R}).$$

Since $g'(\pi/2) = 1 \neq 0$, and f' and g' are both continuous, we deduce, by the Combination Rules for continuous functions, that

$$\begin{aligned}\lim_{x\to\pi/2} \frac{f'(x)}{g'(x)} &= \frac{f'(\pi/2)}{g'(\pi/2)} \\ &= \frac{(-3)\times(-1)}{1} = 3.\end{aligned}$$

Thus, by l'Hôpital's Rule, the required limit exists and

$$\lim_{x\to\pi/2} \frac{\cos 3x}{\sin x - e^{\cos x}} = 3.$$

The next example requires *two* applications of l'Hôpital's Rule.

Worked Exercise F22

Prove that

$$\lim_{x\to 0} \frac{x^2}{\cosh x - 1}$$

exists, and determine its value.

Solution

Let $I = \mathbb{R}$ and define

$$f(x) = x^2 \quad \text{and} \quad g(x) = \cosh x - 1 \quad (x \in \mathbb{R}).$$

Then f and g are differentiable, and $f(0) = g(0) = 0$; hence f and g satisfy the conditions of l'Hôpital's Rule at the point $x = 0$.

Now the derivatives of f and g are

$$f'(x) = 2x \quad \text{and} \quad g'(x) = \sinh x \quad (x \in \mathbb{R}).$$

Thus, by l'Hôpital's Rule, the required limit exists and equals

$$\lim_{x\to 0} \frac{2x}{\sinh x}, \tag{$*$}$$

provided that the limit $(*)$ exists.

Since $g'(0) = 0$, we cannot assert that

$$\lim_{x\to 0} \frac{f'(x)}{g'(x)} = \frac{f'(0)}{g'(0)}.$$

So we try to apply l'Hôpital's Rule a second time.

Both f' and g' are differentiable, and $f'(0) = g'(0) = 0$; hence f' and g' satisfy the conditions of l'Hôpital's Rule at the point $x = 0$. Now

$$f''(x) = 2 \quad \text{and} \quad g''(x) = \cosh x \quad (x \in \mathbb{R}).$$

Thus, by l'Hôpital's Rule, limit $(*)$ exists and equals

$$\lim_{x\to 0} \frac{2}{\cosh x},$$

provided that this limit exists. But the function cosh is continuous on $\mathbb{R}$, and $\cosh 0 = 1$. Thus, by the Quotient Rule for continuous functions,

$$\lim_{x\to 0} \frac{2}{\cosh x} = \frac{2}{1} = 2.$$

Working backwards, we conclude that limit $(*)$ exists and equals 2, so

$$\lim_{x\to 0} \frac{x^2}{\cosh x - 1} = 2.$$

Before asking you to apply l'Hôpital's Rule for yourself, we emphasise that you should *always* check carefully that its conditions hold, because a careless application can easily give an incorrect answer, as you will see in the following worked exercise!

Worked Exercise F23

Explain why the following proof is incorrect and find the correct value for the limit.

Claim (incorrect!)

$$\lim_{x \to 1} \frac{2x^2 - x - 1}{x^2 - x} = 2.$$

Proof (incorrect!) Let $I = \mathbb{R}$ and define

$$f(x) = 2x^2 - x - 1 \quad \text{and} \quad g(x) = x^2 - x \quad (x \in \mathbb{R}).$$

Then f and g are differentiable on $\mathbb{R}$, and

$$f(1) = g(1) = 0.$$

So, by l'Hôpital's Rule,

$$\begin{aligned}\lim_{x \to 1} \frac{f(x)}{g(x)} &= \lim_{x \to 1} \frac{f'(x)}{g'(x)} = \lim_{x \to 1} \frac{4x - 1}{2x - 1} \\ &= \lim_{x \to 1} \frac{f''(x)}{g''(x)} = \lim_{x \to 1} \frac{4}{2} = 2.\end{aligned}$$

■

Solution

The evaluation of the limit in the argument above involves *two* applications of l'Hôpital's Rule. The first application is valid, since f and g are differentiable on $\mathbb{R}$ and $f(1) = g(1) = 0$, so the conditions of l'Hôpital's Rule are satisfied. However, the conditions were not checked for the second application of the Rule, and in fact $f'(1) = 3 \neq 0$ and $g'(1) = 1 \neq 0$, so the conditions are *not* satisfied.

To evaluate the limit correctly, we should instead have concluded that, since f' and g' are both continuous, it follows from the Quotient Rule for continuous functions that

$$\lim_{x \to 1} \frac{f(x)}{g(x)} = \lim_{x \to 1} \frac{f'(x)}{g'(x)} = \frac{f'(1)}{g'(1)} = 3.$$

Exercise F32

Prove that the following limits exist, and evaluate them.

(a) $\displaystyle\lim_{x\to\pi} \frac{\sinh(x-\pi)}{\sin 3x}$ (b) $\displaystyle\lim_{x\to 0} \frac{(1+x)^{1/5}-(1-x)^{1/5}}{(1+2x)^{2/5}-(1-2x)^{2/5}}$

(c) $\displaystyle\lim_{x\to 0} \frac{\sin(x^2)}{1-\cos 4x}$ (d) $\displaystyle\lim_{x\to 0} \frac{\sin x - x\cos x}{x^3}$

Summary

In this unit you have met a formal definition of what it means for a function to be differentiable at a point and seen that, if a function f is differentiable at a point c, then the graph of f has a tangent at the point $(c, f(c))$. You have also seen how to use the definition to obtain the derivatives of some basic functions, and met many rules – in particular, the Glue Rule, the Combination Rules, the Composition Rule and the Inverse Function Rule – that enable us to show that many more functions are differentiable and to determine their derivatives.

You have studied the properties of functions that are differentiable on an interval, and seen that local minima and maxima occur at places where the derivative is zero. You have met Rolle's Theorem and seen how this is used in the proof of the Mean Value Theorem. This has useful corollaries such as the Increasing–Decreasing Theorem, which says that if the derivative of a function is positive then the function is increasing, while if the derivative is negative then the function is decreasing. This in turn enables us to use the second derivative test to determine whether a local extreme value is a maximum or minimum. Finally, you learnt how to use l'Hôpital's Rule to evaluate many limits of the form $\displaystyle\lim_{x\to c} \frac{f(x)}{g(x)}$ where f and g are differentiable functions with $f(c) = g(c) = 0$.

Learning outcomes

After working through this unit, you should be able to:

- explain what is meant by a *differentiable function*, and understand the geometric significance of differentiability
- determine, using the definition, whether or not a function is differentiable at a point
- explain what is meant by a *second derivative* and a higher-order derivative
- explain what is meant by the *left derivative* and the *right derivative* of a function at a given point
- state and use the Glue Rule for differentiation
- use the table of standard derivatives
- use the rules for differentiation to prove the differentiability of a particular function and to calculate its derivative
- state and use the Local Extreme Value Theorem
- state and use Rolle's Theorem
- state and use the Mean Value Theorem
- state and use the Increasing–Decreasing Theorem and the Zero Derivative Theorem
- understand the statements of Cauchy's Mean Value Theorem and l'Hôpital's Rule
- use l'Hôpital's Rule to evaluate certain limits of the form $\lim_{x \to c} \frac{f(x)}{g(x)}$, where $f(c) = g(c) = 0$.

Table of standard derivatives

$f(x)$	$f'(x)$	Domain of f'
k	0	$\mathbb{R}$
x	1	$\mathbb{R}$
$x^n,\ n \in \mathbb{Z} - \{0\}$	nx^{n-1}	$\mathbb{R}$ or $\mathbb{R} - \{0\}$
$x^\alpha,\ \alpha \in \mathbb{R}$	$\alpha x^{\alpha-1}$	$\mathbb{R}^+$
$a^x,\ a > 0$	$a^x \log a$	$\mathbb{R}$
$\sin x$	$\cos x$	$\mathbb{R}$
$\cos x$	$-\sin x$	$\mathbb{R}$
$\tan x$	$\sec^2 x$	$\mathbb{R} - \left\{\left(n + \frac{1}{2}\right)\pi : n \in \mathbb{Z}\right\}$
$\operatorname{cosec} x$	$-\operatorname{cosec} x \cot x$	$\mathbb{R} - \{n\pi : n \in \mathbb{Z}\}$
$\sec x$	$\sec x \tan x$	$\mathbb{R} - \left\{\left(n + \frac{1}{2}\right)\pi : n \in \mathbb{Z}\right\}$
$\cot x$	$-\operatorname{cosec}^2 x$	$\mathbb{R} - \{n\pi : n \in \mathbb{Z}\}$
$\sin^{-1} x$	$1/\sqrt{1 - x^2}$	$(-1, 1)$
$\cos^{-1} x$	$-1/\sqrt{1 - x^2}$	$(-1, 1)$
$\tan^{-1} x$	$1/\left(1 + x^2\right)$	$\mathbb{R}$
e^x	e^x	$\mathbb{R}$
$\log x$	$1/x$	$\mathbb{R}^+$
$\sinh x$	$\cosh x$	$\mathbb{R}$
$\cosh x$	$\sinh x$	$\mathbb{R}$
$\tanh x$	$\operatorname{sech}^2 x$	$\mathbb{R}$
$\sinh^{-1} x$	$1/\sqrt{1 + x^2}$	$\mathbb{R}$
$\cosh^{-1} x$	$1/\sqrt{x^2 - 1}$	$(1, \infty)$
$\tanh^{-1} x$	$1/\left(1 - x^2\right)$	$(-1, 1)$

Solutions to exercises

Solution to Exercise F15

(a) The difference quotient for f at c, where $c \neq 0$, is

$$Q(h) = \frac{f(c+h) - f(c)}{h} = \frac{\frac{1}{c+h} - \frac{1}{c}}{h} = \frac{-1}{(c+h)c}, \quad \text{where } h \neq 0.$$

Thus $Q(h) \to -1/c^2$ as $h \to 0$. Hence f is differentiable at c, with $f'(c) = -1/c^2$.

(b) The difference quotient for f at 0 is

$$Q(h) = \frac{f(h) - f(0)}{h} = \frac{h^2 \cos(1/h) - 0}{h} = h\cos(1/h), \quad \text{where } h \neq 0.$$

Now, $|\cos(1/h)| \leq 1$ for $h \neq 0$, so

$$|Q(h)| \leq |h|, \quad \text{for } h \neq 0.$$

Thus $Q(h) \to 0$ as $h \to 0$, by the Squeeze Rule for limits. Hence f is differentiable at 0, with $f'(0) = 0$.

(c) The difference quotient for f at 0 is

$$Q(h) = \frac{f(h) - f(0)}{h} = \frac{|h| - 0}{h} = \begin{cases} 1, & h > 0, \\ -1, & h < 0. \end{cases}$$

Now consider the two null sequences

$$h_n = \frac{1}{n} \quad \text{and} \quad k_n = -\frac{1}{n}, \quad n = 1, 2, \ldots.$$

These sequences have non-zero terms, and

$$Q(h_n) = 1 \to 1 \ \text{ as } n \to \infty,$$

but

$$Q(k_n) = -1 \to -1 \ \text{ as } n \to \infty.$$

Since these limits are different, f is not differentiable at 0.

(d) The difference quotient for f at 0 is

$$Q(h) = \frac{f(h) - f(0)}{h} = \frac{|h|^{1/2}\sin(1/h) - 0}{h} = \begin{cases} \dfrac{\sin(1/h)}{h^{1/2}}, & h > 0, \\ -\dfrac{\sin(1/h)}{|h|^{1/2}}, & h < 0. \end{cases}$$

Now consider the null sequence

$$h_n = \frac{1}{\left(2n + \frac{1}{2}\right)\pi}, \quad n = 1, 2, \ldots,$$

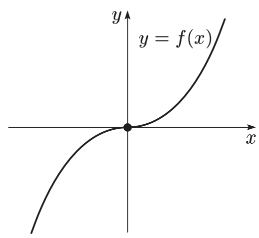
which has positive terms. This gives

$$Q(h_n) = \frac{\sin(1/h_n)}{h_n^{1/2}} = \left(2n + \tfrac{1}{2}\right)^{1/2} \pi^{1/2} \sin\left(2n + \tfrac{1}{2}\right)\pi = \left(2n + \tfrac{1}{2}\right)^{1/2} \pi^{1/2} \to \infty \ \text{ as } n \to \infty.$$

Hence f is not differentiable at 0.

Solution to Exercise F16

The graph of $y = f(x)$ is given below.



(This is included to aid your understanding – you do not need to sketch the graph as part of your solution.)

Let $I = \mathbb{R}$ and define

$$g(x) = -x^2 \quad (x \in \mathbb{R}) \quad \text{and} \quad h(x) = x^2 \quad (x \in \mathbb{R}).$$

Then

$$\begin{aligned} f(x) &= g(x), \quad \text{for } x < 0, \\ f(x) &= h(x), \quad \text{for } x > 0, \end{aligned} \qquad (*)$$

so condition 1 of the Glue Rule holds (with $c = 0$).

Furthermore, $f(0) = g(0) = h(0) = 0$, so condition 2 holds, and g and h are differentiable with

$$g'(x) = -2x \quad (x \in \mathbb{R}) \quad \text{and} \quad h'(x) = 2x \quad (x \in \mathbb{R}),$$

so condition 3 holds.

Since $g'(0) = h'(0) = 0$, it follows from the Glue Rule that f is differentiable at 0 and $f'(0) = 0$.

Also, by equations $(*)$,

$$\begin{aligned} f'(x) &= g'(x) = -2x, \quad \text{for } x < 0, \\ f'(x) &= h'(x) = 2x, \quad \text{for } x > 0, \end{aligned}$$

since differentiability is a local property.

Hence f is differentiable (on $\mathbb{R}$), and

$$f'(x) = \begin{cases} -2x, & x < 0, \\ 0, & x = 0, \\ 2x, & x > 0. \end{cases}$$

Thus

$$f'(x) = 2|x| \quad (x \in \mathbb{R}).$$

Solution to Exercise F17

In each case we use the Combination Rules.

(a) $f'(x) = 7x^6 - 8x^3 + 9x^2 - 5 \quad (x \in \mathbb{R})$

(b) $$\begin{aligned} f'(x) &= \frac{(x^3-1)2x - (x^2+1)3x^2}{(x^3-1)^2} \\ &= \frac{-x^4 - 3x^2 - 2x}{(x^3-1)^2} \quad (x \in \mathbb{R} - \{1\}) \end{aligned}$$

(c) $$\begin{aligned} f'(x) &= \cos^2 x - \sin^2 x \\ &= \cos 2x \quad (x \in \mathbb{R}) \end{aligned}$$

(d) $$\begin{aligned} f'(x) \\ &= \frac{(3 + \sin x - 2\cos x)e^x - e^x(\cos x + 2\sin x)}{(3 + \sin x - 2\cos x)^2} \\ &= \frac{e^x(3 - \sin x - 3\cos x)}{(3 + \sin x - 2\cos x)^2} \quad (x \in \mathbb{R}) \end{aligned}$$

Solution to Exercise F18

We have

$$\begin{aligned} f'(x) &= e^x + xe^x = e^x(1+x), \\ f''(x) &= e^x(1+x) + e^x = e^x(2+x), \\ f^{(3)}(x) &= e^x(2+x) + e^x = e^x(3+x). \end{aligned}$$

Solution to Exercise F19

In each case, we use the Quotient Rule and the derivatives of sin and cos.

(a) $f(x) = \tan x = \sin x / \cos x$, so

$$\begin{aligned} f'(x) &= \frac{\cos x \cos x - \sin x(-\sin x)}{\cos^2 x} \\ &= \frac{1}{\cos^2 x} = \sec^2 x \end{aligned}$$

on the domain of f.

(b) $f(x) = \operatorname{cosec} x = 1/\sin x$, so

$$\begin{aligned} f'(x) &= -\frac{\cos x}{\sin^2 x} \\ &= -\operatorname{cosec} x \cot x \end{aligned}$$

on the domain of f.

(c) $f(x) = \sec x = 1/\cos x$, so

$$\begin{aligned} f'(x) &= \frac{\sin x}{\cos^2 x} \\ &= \sec x \tan x \end{aligned}$$

on the domain of f.

(d) $f(x) = \cot x = \cos x / \sin x$, so

$$\begin{aligned} f'(x) &= \frac{\sin x(-\sin x) - \cos x \cos x}{\sin^2 x} \\ &= -\frac{1}{\sin^2 x} = -\operatorname{cosec}^2 x \end{aligned}$$

on the domain of f.

Solution to Exercise F20

In each case we use the Combination Rules.

(a) $f(x) = \sinh x = \frac{1}{2}(e^x - e^{-x})$, so

$$\begin{aligned} f'(x) &= \tfrac{1}{2}(e^x + e^{-x}) \\ &= \cosh x. \end{aligned}$$

(b) $f(x) = \cosh x = \frac{1}{2}(e^x + e^{-x})$, so

$$\begin{aligned} f'(x) &= \tfrac{1}{2}(e^x - e^{-x}) \\ &= \sinh x. \end{aligned}$$

(c) $f(x) = \tanh x = \sinh x / \cosh x$, so

$$\begin{aligned} f'(x) &= \frac{\cosh x \cosh x - \sinh x \sinh x}{\cosh^2 x} \\ &= \frac{1}{\cosh^2 x} = \operatorname{sech}^2 x. \end{aligned}$$

Solution to Exercise F21

(a) $f(x) = \sinh(x^2)$, so

$$f'(x) = 2x\cosh(x^2).$$

(b) $f(x) = \sin(\sinh 2x)$, so

$$f'(x) = 2\cos(\sinh 2x)\cosh 2x.$$

(c) $f(x) = \sin\left(\dfrac{\cos 2x}{x^2}\right) \quad (x \in (0,\infty))$,

so on this interval

$$\begin{aligned} f'(x) &= \cos\left(\frac{\cos 2x}{x^2}\right)\left(\frac{x^2(-2\sin 2x) - 2x\cos 2x}{x^4}\right) \\ &= -\frac{2}{x^3}(x\sin 2x + \cos 2x)\cos\left(\frac{\cos 2x}{x^2}\right). \end{aligned}$$

Solution to Exercise F22

(a) The function

$$f(x) = \cos x \quad (x \in (0,\pi))$$

is continuous and strictly decreasing, and

$$f((0,\pi)) = (-1,1).$$

Also, f is differentiable on $(0,\pi)$, and its derivative $f'(x) = -\sin x$ is non-zero there.

Thus f satisfies the conditions of the Inverse Function Rule.

Hence f has an inverse function f^{-1} and $f^{-1} = \cos^{-1}$ is differentiable on $(-1,1)$. If $y = f(x) = \cos x$, then

$$(f^{-1})'(y) = \frac{1}{f'(x)} = -\frac{1}{\sin x}.$$

Since $\sin x > 0$ on $(0,\pi)$ and $\sin^2 x + \cos^2 x = 1$, it follows that

$$\sin x = \sqrt{1-\cos^2 x} = \sqrt{1-y^2},$$

so

$$(f^{-1})'(y) = \frac{-1}{\sqrt{1-y^2}}.$$

Replacing the domain variable y by x, we obtain

$$(\cos^{-1})'(x) = \frac{-1}{\sqrt{1-x^2}} \quad (x \in (-1,1)).$$

(b) The function

$$f(x) = \sinh x \quad (x \in \mathbb{R})$$

is continuous and strictly increasing, and $f(\mathbb{R}) = \mathbb{R}$.

Also, f is differentiable on $\mathbb{R}$, and its derivative $f'(x) = \cosh x$ is non-zero there.

Thus f satisfies the conditions of the Inverse Function Rule.

Hence f has an inverse function f^{-1} and $f^{-1} = \sinh^{-1}$ is differentiable on $\mathbb{R}$. If $y = f(x) = \sinh x$, then

$$(f^{-1})'(y) = \frac{1}{f'(x)} = \frac{1}{\cosh x}.$$

Since $\cosh x > 0$ on $\mathbb{R}$ and $\cosh^2 x = 1 + \sinh^2 x$, it follows that

$$\cosh x = \sqrt{1+\sinh^2 x} = \sqrt{1+y^2},$$

so

$$(f^{-1})'(y) = \frac{1}{\sqrt{1+y^2}}.$$

Replacing the domain variable y by x, we obtain

$$(\sinh^{-1})'(x) = \frac{1}{\sqrt{1+x^2}} \quad (x \in \mathbb{R}).$$

Solution to Exercise F23

(a) If $x_1 < x_2$, then $x_1^5 < x_2^5$, so $f(x_1) < f(x_2)$. Thus f is strictly increasing and continuous on $\mathbb{R}$, and $f(\mathbb{R}) = \mathbb{R}$. Also

$$f'(x) = 5x^4 + 1 \neq 0, \quad \text{for } x \in \mathbb{R}.$$

Thus f satisfies the conditions of the Inverse Function Rule. Hence f has an inverse function f^{-1} which is differentiable on $\mathbb{R}$.

(b) Now, $f(0) = -1$, $f(1) = 1$ and $f(-1) = -3$. Hence, by the Inverse Function Rule,

$$(f^{-1})'(-1) = \frac{1}{f'(0)} = 1,$$

$$(f^{-1})'(1) = \frac{1}{f'(1)} = \tfrac{1}{6},$$

$$(f^{-1})'(-3) = \frac{1}{f'(-1)} = \tfrac{1}{6}.$$

Solution to Exercise F24

By definition,

$$f(x) = x^x = \exp(x \log x) \quad (x \in \mathbb{R}^+).$$

The functions $x \longmapsto x$ and $x \longmapsto \log x$ are differentiable on $\mathbb{R}^+$, and exp is differentiable on $\mathbb{R}$. It follows by the Product Rule and the Composition Rule that f is differentiable on $\mathbb{R}^+$, and that

$$\begin{aligned} f'(x) &= \exp(x \log x)\,(\log x + x \times (1/x)) \\ &= x^x(\log x + 1) \quad (x \in \mathbb{R}^+). \end{aligned}$$

Solution to Exercise F25

Since the functions sin and cos are continuous and differentiable on $\mathbb{R}$, so also is f, by the Combination Rules.

Now,

$$f'(x) = 2\sin x \cos x - \sin x = \sin x(2\cos x - 1);$$

thus f' vanishes on $(0, \pi/2)$ when $\cos x = \frac{1}{2}$, that is, when $x = \pi/3$.

Since $f(0) = 1$, $f(\pi/2) = 1$ and

$$f(\pi/3) = (\sqrt{3}/2)^2 + \tfrac{1}{2} = \tfrac{3}{4} + \tfrac{1}{2} = \tfrac{5}{4},$$

it follows that on $[0, \pi/2]$:

the minimum of f is 1, occurring at $x = 0$ and $\pi/2$;
the maximum of f is $\frac{5}{4}$, occurring at $x = \pi/3$.

Solution to Exercise F26

Since f is a polynomial function, f is continuous on $[1, 3]$ and differentiable on $(1, 3)$. Also, $f(1) = 2$ and $f(3) = 2$, so $f(1) = f(3)$.

Thus f satisfies the conditions of Rolle's Theorem on $[1, 3]$, so there exists c in $(1, 3)$ such that $f'(c) = 0$.

(In fact, since

$$f'(x) = 4x^3 - 12x^2 + 6x = 2x(2x^2 - 6x + 3),$$

and $2x^2 - 6x + 3 = 0$ for $x = \frac{1}{2}(3 \pm \sqrt{3})$, we have $c = \frac{1}{2}(3 + \sqrt{3}) \simeq 2.37$.)

Solution to Exercise F27

(a) No: f is not defined at $\pi/2$.

(b) No: f is not differentiable at 1.

(c) Yes: all the conditions are satisfied.

(d) No: $f(0) \neq f(\pi/2)$.

Solution to Exercise F28

The function $f(x) = xe^x$ is continuous on $[0, 2]$ and differentiable on $(0, 2)$ by the Product Rule. Thus f satisfies the conditions of the Mean Value Theorem on $[0, 2]$.

Now,

$$\frac{f(2) - f(0)}{2 - 0} = \frac{2e^2 - 0}{2} = e^2.$$

Thus, by the Mean Value Theorem, there exists a point c in $(0, 2)$ such that $f'(c) = e^2$.

Solution to Exercise F29

(a) The function f is continuous on I and differentiable on the interior of I and so we can apply the Increasing–Decreasing Theorem. We have $f'(x) = 4x^{1/3} - 4 = 4(x^{1/3} - 1)$. Thus $f'(x) > 0$ for $x \in (1, \infty)$, so f is strictly increasing on $[1, \infty)$, by the strict inequalities version of the Increasing–Decreasing Theorem.

(b) The function f is continuous on I and differentiable on the interior of I and so we can apply the Increasing–Decreasing Theorem. We have $f'(x) = 1 - 1/x = (x - 1)/x$. Thus $f'(x) < 0$ for $x \in (0, 1)$, so f is strictly decreasing on $(0, 1]$, by the strict inequalities version of the Increasing–Decreasing Theorem.

Solution to Exercise F30

(a) We have

$$f'(x) = 3x^2 - 6x = 3x(x - 2).$$

Thus $f'(x) = 0$ for $x = 0$ and 2, so the required values of c are 0 and 2.

(b) We have

$$f''(x) = 6x - 6,$$

so

$$f''(0) = -6 < 0 \quad \text{and} \quad f''(2) = 6 > 0.$$

Also, $f(0) = 1$ and $f(2) = -3$.

Since f is a twice-differentiable function defined on $\mathbb{R}$ and f'' is continuous on $\mathbb{R}$, it follows from the Second Derivative Test that f has a local maximum of 1 at $x = 0$ and a local minimum of -3 at $x = 2$.

Solution to Exercise F31

In each case we follow the steps in Strategy F7.

(a) 1. Let

$$f(x) = x - \sin x \quad (x \in [0, \infty)).$$

Then f is continuous on $[0, \infty)$ and differentiable on $(0, \infty)$.

2. We have

$$f'(x) = 1 - \cos x \geq 0, \quad \text{for } x \in (0, \infty),$$

and $f(0) = 0$.

Thus f is increasing on $[0, \infty)$, by the Increasing–Decreasing Theorem, so

$$f(x) \geq f(0) = 0, \quad \text{for } x \in [0, \infty).$$

Hence

$$\sin x \leq x, \quad \text{for } x \in [0, \infty).$$

(b) 1. Let

$$f(x) = \tfrac{2}{3}x + \tfrac{1}{3} - x^{2/3} \quad (x \in [0, 1]).$$

Then f is continuous on $[0, 1]$ and differentiable on $(0, 1)$.

2. We have

$$\begin{aligned} f'(x) &= \tfrac{2}{3} - \tfrac{2}{3}x^{-1/3} \\ &= \tfrac{2}{3}(1 - x^{-1/3}) < 0, \quad \text{for } x \in (0, 1), \end{aligned}$$

and $f(1) = \frac{2}{3} + \frac{1}{3} - 1 = 0$.

Thus f is decreasing on $[0, 1]$, by the Increasing–Decreasing Theorem, so

$$f(x) \geq f(1) = 0, \quad \text{for } x \in [0, 1].$$

Hence

$$\tfrac{2}{3}x + \tfrac{1}{3} \geq x^{2/3}, \quad \text{for } x \in [0, 1].$$

Solution to Exercise F32

(a) Let $I = \mathbb{R}$ and define

$$f(x) = \sinh(x - \pi) \quad (x \in \mathbb{R})$$

and

$$g(x) = \sin 3x \quad (x \in \mathbb{R}).$$

Then f and g are differentiable on $\mathbb{R}$, and

$$f(\pi) = g(\pi) = 0.$$

Thus f and g satisfy the conditions of l'Hôpital's Rule at the point $x = 0$.

Now,

$$f'(x) = \cosh(x - \pi) \quad \text{and} \quad g'(x) = 3\cos 3x.$$

Since $g'(\pi) = -3 \neq 0$, and f' and g' are both continuous, we deduce, by the Combination Rules for continuous functions, that

$$\lim_{x \to \pi} \frac{f'(x)}{g'(x)} = \frac{f'(\pi)}{g'(\pi)} = \frac{1}{-3} = -\tfrac{1}{3}.$$

Thus, by l'Hôpital's Rule, the required limit exists and equals $-\frac{1}{3}$.

(b) Let $I = \left(-\frac{1}{2}, \frac{1}{2}\right)$, say, and define

$$f(x) = (1 + x)^{1/5} - (1 - x)^{1/5}$$

and

$$g(x) = (1 + 2x)^{2/5} - (1 - 2x)^{2/5}$$

on I. (The only requirements when selecting the open interval I are that it must contain 0 and lie in the domains of both f and g; that is, all $x \in I$ must satisfy $1 - 2x \geq 0$ and $1 + 2x \geq 0$.)

Then f and g are differentiable on I and

$$f(0) = g(0) = 0.$$

Thus f and g satisfy the conditions of l'Hôpital's Rule at the point $x = 0$.

Now,

$$f'(x) = \tfrac{1}{5}(1 + x)^{-4/5} + \tfrac{1}{5}(1 - x)^{-4/5}$$

and

$$g'(x) = \tfrac{4}{5}(1 + 2x)^{-3/5} + \tfrac{4}{5}(1 - 2x)^{-3/5}.$$

Since $g'(0) = \frac{4}{5} + \frac{4}{5} \neq 0$, and f' and g' are both continuous, we deduce, by the Combination Rules for continuous functions, that

$$\lim_{x\to 0} \frac{f'(x)}{g(x)} = \frac{f'(0)}{g'(0)} = \frac{\frac{1}{5}+\frac{1}{5}}{\frac{4}{5}+\frac{4}{5}} = \tfrac{1}{4}.$$

Thus, by l'Hôpital's Rule, the required limit exists and equals $\frac{1}{4}$.

(c) Let $I = \mathbb{R}$ and define

$$f(x) = \sin(x^2) \quad \text{and} \quad g(x) = 1 - \cos 4x \quad (x \in \mathbb{R}).$$

Then f and g are differentiable on $\mathbb{R}$, and

$$f(0) = g(0) = 0.$$

Thus f and g satisfy the conditions of l'Hôpital's Rule at the point $x = 0$.

Now,

$$f'(x) = 2x\cos(x^2) \quad \text{and} \quad g'(x) = 4\sin 4x.$$

Thus, by l'Hôpital's Rule, the required limit exists and equals

$$\lim_{x\to 0} \frac{f'(x)}{g'(x)}, \qquad (*1)$$

provided that limit $(*1)$ exists. Here $f'(0) = g'(0) = 0$, so we cannot apply l'Hôpital's Rule at this stage. However, f' and g' are differentiable on $\mathbb{R}$, so f' and g' satisfy the conditions of l'Hôpital's Rule at the point $x = 0$.

Now,

$$f''(x) = -(2x)^2\sin(x^2) + 2\cos(x^2)$$

and

$$g''(x) = 16\cos 4x.$$

Thus, by l'Hôpital's Rule, the required limit exists and equals

$$\lim_{x\to 0} \frac{f''(x)}{g''(x)}, \qquad (*2)$$

provided that limit $(*2)$ exists.

Since $g''(0) = 16 \neq 0$, and f'' and g'' are both continuous, we deduce, by the Combination Rules for continuous functions, that

$$\lim_{x\to 0} \frac{f''(x)}{g''(x)} = \frac{f''(0)}{g''(0)} = \frac{2}{16} = \tfrac{1}{8}.$$

Hence limit $(*2)$ exists and equals $\frac{1}{8}$.

Thus limit $(*1)$ exists and equals $\frac{1}{8}$, so the required limit also exists and equals $\frac{1}{8}$.

(d) Let $I = \mathbb{R}$ and define

$$f(x) = \sin x - x\cos x \quad \text{and} \quad g(x) = x^3 \quad (x \in \mathbb{R}).$$

Then f and g are differentiable on $\mathbb{R}$, and $f(0) = g(0) = 0$. Thus f and g satisfy the conditions of l'Hôpital's Rule at the point $x = 0$.

Now,

$$f'(x) = x\sin x \quad \text{and} \quad g'(x) = 3x^2.$$

Thus, by l'Hôpital's Rule, the required limit exists and equals

$$\lim_{x\to 0} \frac{f'(x)}{g'(x)},$$

provided that this limit exists. But

$$\lim_{x\to 0} \frac{f'(x)}{g'(x)} = \lim_{x\to 0} \frac{x\sin x}{3x^2} = \frac{1}{3}\lim_{x\to 0} \frac{\sin x}{x} = \frac{1}{3}.$$

Hence, the required limit exists and equals $\frac{1}{3}$.

Unit F3

Integration

Introduction

In this unit you will study the question:

> What do we mean by the area between the graph of a real function f and the x-axis?

You will see how this can be answered by trapping the required area between increasingly accurate lower and upper estimates, each of which is the sum of the areas of suitably chosen rectangles. The area between the graph $y = f(x)$ and the segment $[a, b]$ of the x-axis is *defined* to be the supremum of the lower estimates and the infimum of the upper estimates, as long as these two values are equal. In this case, we call the common value the *integral* of f on $[a, b]$, written

$$\int_a^b f \quad \text{or} \quad \int_a^b f(x)\,dx.$$

You will see that, for many functions, we can evaluate integrals by using the *Fundamental Theorem of Calculus*, which allows us to think of integration as the inverse operation of differentiation. Although we will review techniques of integration, our main focus in this unit is on providing a rigorous foundation for the idea of integration, and on showing how this relates to concepts you have met in previous analysis units.

Often it is not possible to evaluate an integral explicitly, and later in the unit you will meet methods for obtaining upper and lower bounds for the integral in such cases. You will also see how we can apply integration to derive some remarkable formulas for π and for estimating factorials, and to give a useful additional test for the convergence of certain series.

1 The Riemann integral

The purpose of this section is to give a rigorous definition of what we mean by the area between the graph

$$y = f(x) \quad (x \in [a, b])$$

and the closed interval $[a, b]$ on the x-axis, and to explore its implications. We begin with an informal discussion to set the scene.

We have an intuitive notion of area. There are formulas for calculating the areas of simple geometric shapes such as rectangles and triangles, and we would certainly want our rigorous definition to agree with these. We would probably also agree that the region between the x-axis and the graph of a *continuous* function defined on a closed interval $[a, b]$ has a definite area, even if we are uncertain how to measure it. On the other hand, for discontinuous functions it is not obvious that we can always say that the region between a graph and the x-axis *has* an area. For example, can we define such an area for the function

$$f(x) = \begin{cases} x^2, & 0 \leq x \leq 1, \\ 2, & 1 < x \leq 2, \end{cases}$$

which has a discontinuity at the point $x = 1$? The graph of f is illustrated in Figure 1(a).

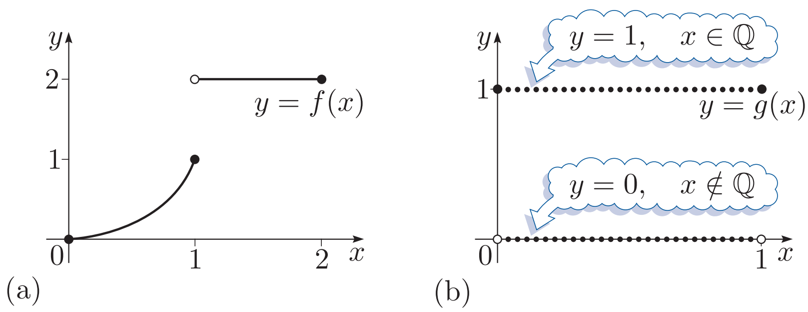


Figure 1 The graphs of the discontinuous functions f and g

As another example, can we define the area between the x-axis and the graph of the **Dirichlet function** on the closed interval $[0, 1]$, illustrated in Figure 1(b)? You met the Dirichlet function in Subsection 3.2 of Unit F1 *Limits*. Its rule is

$$g(x) = \begin{cases} 1, & 0 \leq x \leq 1,\ x \text{ rational}, \\ 0, & 0 \leq x \leq 1,\ x \text{ irrational}, \end{cases}$$

and in Unit F1 it was shown to be discontinuous at every point of its domain.

It seems desirable that our definition of area should cover a wide range of functions. Later in this section we will prove that we can always assign a value to the area between the graph and the x-axis for a continuous function defined on a closed interval. You will also see that we can do the same for the function illustrated in Figure 1(a), but *not* for the function illustrated in Figure 1(b).

Our definition of area is based on finding lower and upper estimates for the 'area' (if it exists) of the region between the graph $y = f(x)$ and the x-axis, using the following approach. First we divide the interval $[a, b]$ into a set of subintervals, called a *partition* of $[a, b]$. Then we consider two sets of rectangles, each rectangle having one of the subintervals as its base. In one set, we choose rectangles whose top edges lie on or below the graph, so

the sum of their individual areas forms a lower estimate for the 'area' of the region; see Figure 2(a). In the other, we choose rectangles whose top edges lie on or above the graph, so the sum of their individual areas forms an upper estimate for the 'area'; see Figure 2(b).

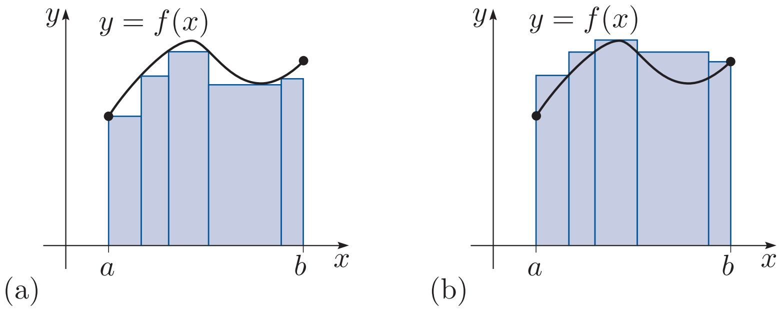


Figure 2 (a) A lower estimate and (b) an upper estimate for the area of the region between the graph $y = f(x)$ and the x-axis

In this way we can obtain many lower estimates and upper estimates by choosing different partitions of $[a, b]$. If there is a real number A with the properties

$$\text{the supremum of the lower estimates} = A$$

and

$$\text{the infimum of the upper estimates} = A,$$

then we define A to be the *area* between the graph and the x-axis. We call the number A the *integral* of f on $[a, b]$, and denote it by

$$\int_a^b f \quad \text{or} \quad \int_a^b f(x)\,dx.$$

We make all these ideas precise in the rest of this section, which is the longest and hardest section of the unit. On a first reading, you may wish to try to understand the main ideas without following every detail. The details may be easier to understand on a second reading.

1.1 Definition of the integral

In this subsection we work towards giving a rigorous definition of the area between the graph of a function f defined on a closed interval $[a, b]$ and the x-axis; that is, the integral of f on $[a, b]$.

Before we can give the definition, we need to introduce a number of key ideas. In the paragraphs below you will study:

- some important terminology for functions
- what is meant by a partition of a closed interval
- the use of lower and upper Riemann sums to estimate areas.

Terminology for functions

In Unit D1 *Numbers* you met the definitions of lower bound, greatest lower bound, upper bound and least upper bound of *sets* in $\mathbb{R}$. Here you will meet analogous definitions for *functions* defined on an interval in $\mathbb{R}$. These definitions of greatest lower bound and least upper bound generalise the notion of the *minimum* and *maximum* of a function which you met in Unit D4 *Continuity*. The minimum and maximum of a function are illustrated in Figure 3 and we give a reminder of their definitions (in a slightly different form from those you saw in Unit D4) together with the definitions of lower and upper bounds for functions.

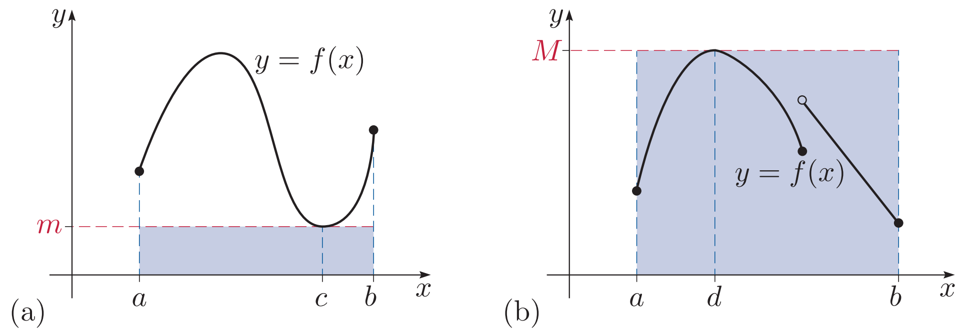


Figure 3 (a) The minimum of a function (b) The maximum of a different function

Definitions

Let the function f be defined on the closed interval $[a, b]$. Then the following hold on $[a, b]$.

- f is **bounded below** on $[a, b]$ with m as a **lower bound** if

  $$f(x) \geq m, \quad \text{for all } x \in [a, b].$$

- m is the **minimum** of f on $[a, b]$ if
  1. m is a lower bound for f on $[a, b]$, and
  2. $f(c) = m$, for some $c \in [a, b]$.

  Thus $m = \min\{f(x) : a \leq x \leq b\}$, which we also write as $\min\limits_{[a,b]} f$ or simply as $\min f$.

- f is **bounded above** on $[a, b]$ with M as an **upper bound** if

  $$f(x) \leq M, \quad \text{for all } x \in [a, b].$$

- M is the **maximum** of f on $[a, b]$ if
  1. M is an upper bound for f on $[a, b]$, and
  2. $f(d) = M$, for some $d \in [a, b]$.

  Thus $M = \max\{f(x) : a \leq x \leq b\}$, which we also write as $\max\limits_{[a,b]} f$ or simply as $\max f$.

- f is **bounded** on $[a, b]$ if it is both bounded below and bounded above on $[a, b]$.

Note that any lower bound for a function f on $[a, b]$ is less than or equal to any upper bound for f on $[a, b]$.

A function f that is continuous on a closed interval $[a, b]$ necessarily has both a minimum and a maximum, by the Extreme Value Theorem for continuous functions, as you saw in Subsection 3.3 of Unit D4. However, if f is not continuous on $[a, b]$, then it may or may not have a minimum or a maximum on $[a, b]$; for example, the function in Figure 4(a) has neither a minimum nor a maximum on $[a, b]$, whilst the function in Figure 4(b) has a maximum but no minimum on $[a, b]$.

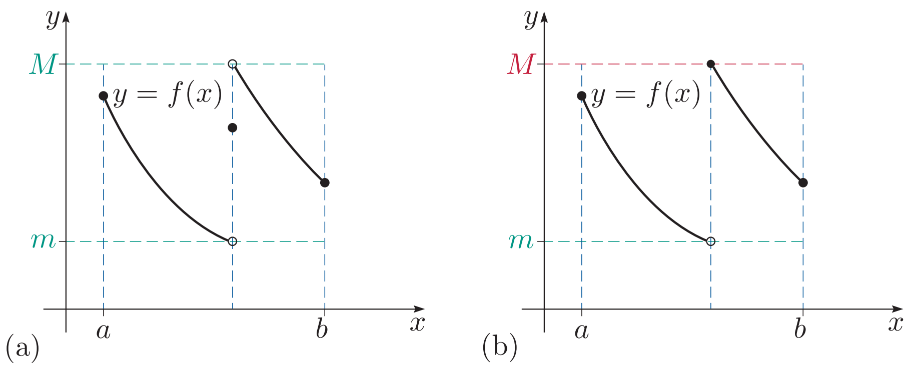


Figure 4 Two bounded functions on $[a, b]$: (a) a function with neither a minimum nor a maximum (b) a function with a maximum but no minimum

Both functions in Figure 4 are certainly bounded on $[a, b]$; and there are numbers m and M, as shown, such that $m \leq f(x) \leq M$ for all $x \in [a, b]$. However, although the first function f takes values as close as we please to m and M, there is no point x in $[a, b]$ where $f(x) = m$ or M; and although the second function f takes values as close as we please to m, there is no point x in $[a, b]$ where $f(x) = m$. This suggests the notions of greatest lower bound and least upper bound.

Definitions

Let the function f be defined on the closed interval $[a, b]$. Then

- m is the **infimum** or **greatest lower bound** of f on $[a, b]$ if

  1. m is a lower bound for f on $[a, b]$, and
  2. if $m' > m$, then $f(c) < m'$, for some $c \in [a, b]$.

  Thus $m = \inf\{f(x) : a \leq x \leq b\}$, which we also write as $\inf\limits_{[a,b]} f$ or simply as $\inf f$.

- M is the **supremum** or **least upper bound** of f on $[a, b]$ if

  1. M is an upper bound for f on $[a, b]$, and
  2. if $M' < M$, then $f(d) > M'$, for some $d \in [a, b]$.

  Thus $M = \sup\{f(x) : a \leq x \leq b\}$, which we also write as $\sup\limits_{[a,b]} f$ or simply as $\sup f$.

Remarks

1. Note from the definition that the concepts of the infimum of a real *function* and the infimum of a *set* of real numbers are related in the following way: the infimum of a function is the infimum of the image set of the function. A similar remark applies to supremums.
2. Any function f that is bounded on $[a, b]$ necessarily possesses an infimum and a supremum on $[a, b]$: if m is any lower bound of f on $[a, b]$, then $\inf f \geq m$ and if M is any upper bound of f on $[a, b]$, then $\sup f \leq M$.
3. If $\min\limits_{[a,b]} f$ exists, then $\inf\limits_{[a,b]} f = \min\limits_{[a,b]} f$. Similarly, if $\max\limits_{[a,b]} f$ exists, then $\sup\limits_{[a,b]} f = \max\limits_{[a,b]} f$.

Worked Exercise F24

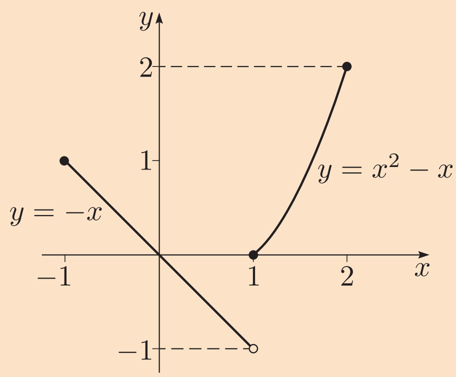
Consider the function

$$f(x) = \begin{cases} -x, & -1 \leq x < 1, \\ x^2 - x, & 1 \leq x \leq 2. \end{cases}$$

Sketch the graph of f and identify $\min f$, $\max f$, $\inf f$ and $\sup f$ (if they exist).

Solution

The graph of f is shown below.



The graph suggests that $\inf f$ is equal to -1 but that $\min f$ does not exist because f does not take the value -1. We begin by checking whether this is true.

First, $\inf f = -1$, since

1. $f(x) \geq -1$, for all $x \in [-1, 2]$,
2. if $m' > -1$, then m' is not a lower bound for f on $[-1, 2]$ because the sequence $(1 - 1/n)$ is contained in $[-1, 2]$ and
$$f(1 - 1/n) = -1 + 1/n \to -1 \text{ as } n \to \infty,$$
so there exists $x' \in [-1, 2]$ such that $f(x') < m'$.

Recall that, if $\min f$ exists, then it is equal to $\inf f$.

Also, $\min f$ does not exist since there is no point x such that $f(x) = -1$.

Next, $\max f = 2$, since

1. $f(x) \leq 2$, for all $x \in [-1, 2]$,
2. $f(2) = 2$.

Finally, $\sup f = 2$, since f has maximum 2 on $[-1, 2]$.

Exercise F33

For each of the following functions f on $[-1, 1]$, sketch the graph of f and identify $\min f$, $\max f$, $\inf f$ and $\sup f$ (if they exist).

(a) $f(x) = \begin{cases} x^2, & -1 < x < 1, \\ \frac{1}{2}, & x = \pm 1. \end{cases}$

(b) $f(x) = \begin{cases} x^2, & -1 \leq x \leq 0, \\ x^2 - 1, & 0 < x \leq 1. \end{cases}$

Partitions of a closed interval

We now introduce the notion of a *partition* of a closed interval, which will play an important role in estimating the area between the graph of a function and the x-axis.

Definitions

A **partition** P of a closed interval $[a, b]$ is a collection of a finite number of closed subintervals of $[a, b]$,

$$P = \{[x_0, x_1], [x_1, x_2], \dots, [x_{i-1}, x_i], \dots, [x_{n-1}, x_n]\},$$

where

$$a = x_0 < x_1 < \dots < x_i < \dots < x_n = b.$$

The points x_i, $0 \le i \le n$, are called the **partition points** of P.

The ith **subinterval** is $[x_{i-1}, x_i]$, $1 \le i \le n$, and its **length** is denoted by $\delta x_i = x_i - x_{i-1}$.

The **mesh** of P is the quantity $\|P\| = \max\limits_{1 \le i \le n} \{\delta x_i\}$.

A **standard partition** is a partition with subintervals of equal length.

These definitions are illustrated in Figure 5.

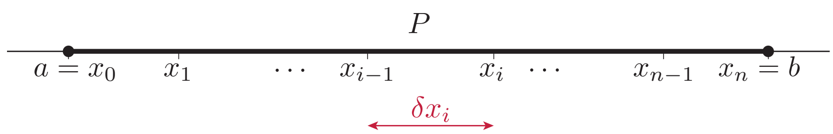


Figure 5 A partition P of an interval $[a, b]$

Worked Exercise F25

Let P be the partition of [0,1] given by

$$P = \left\{ \left[0, \tfrac{1}{2}\right], \left[\tfrac{1}{2}, \tfrac{3}{5}\right], \left[\tfrac{3}{5}, \tfrac{3}{4}\right], \left[\tfrac{3}{4}, 1\right] \right\}.$$

Find the mesh of P.

Solution

The partition points are 0, $\frac{1}{2}$, $\frac{3}{5}$, $\frac{3}{4}$ and 1.

We have

$$\delta x_1 = \tfrac{1}{2} - 0 = \tfrac{1}{2}, \ \delta x_2 = \tfrac{3}{5} - \tfrac{1}{2} = \tfrac{1}{10}, \ \delta x_3 = \tfrac{3}{4} - \tfrac{3}{5} = \tfrac{3}{20},$$
$$\delta x_4 = 1 - \tfrac{3}{4} = \tfrac{1}{4}.$$

The mesh of P is the length of the largest subinterval.

So the mesh of P is

$$\|P\| = \max\left\{ \tfrac{1}{2}, \tfrac{1}{10}, \tfrac{3}{20}, \tfrac{1}{4} \right\} = \tfrac{1}{2}.$$

P is not a standard partition of $[0, 1]$, since not all its subintervals are of equal length.

Exercise F34

Write down the standard partition P of $[-1,2]$ that contains four subintervals, and state the mesh of P.

Lower and upper Riemann sums

Next, we introduce the *lower* and *upper Riemann sums* for a bounded function f on an interval $[a,b]$ with partition P; these correspond to underestimates and overestimates in our intuitive notion of the area between the graph $y=f(x)$ and the x-axis.

Definitions

Let f be a bounded function on $[a,b]$, and let P be the partition $\{[x_0,x_1],[x_1,x_2],\ldots,[x_{i-1},x_i],\ldots,[x_{n-1},x_n]\}$, where $x_0=a$ and $x_n=b$. Let

$$m_i=\inf\{f(x):x\in[x_{i-1},x_i]\}$$

and

$$M_i=\sup\{f(x):x\in[x_{i-1},x_i]\},$$

for $i=1,2,\ldots,n$.

Then the **lower Riemann sum** for f on $[a,b]$ with partition P is

$$L(f,P)=\sum_{i=1}^{n} m_i\delta x_i,$$

and the **upper Riemann sum** for f on $[a,b]$ with partition P is

$$U(f,P)=\sum_{i=1}^{n} M_i\delta x_i.$$

Note that the above definitions work equally well whether f takes positive or negative values in $[a,b]$. Regions between the graph $y=f(x)$ and the x-axis where f takes negative values make a negative contribution to the lower and upper Riemann sums. In this subsection we will in general illustrate results for functions that are *non-negative* throughout $[a,b]$, but we give further consideration to functions that are negative on all or part of the interval $[a,b]$ in Subsection 1.3.

The terms m_i and M_i in the definitions denote the greatest lower bound and least upper bound of f on the ith subinterval of the partition; we need to use the infimum and supremum on the subintervals since the function f may not be continuous and so may not have a minimum or maximum on all (or any) subintervals. The lower Riemann sum is the sum of the areas of the rectangles with height m_i and width δx_i, giving a *lower* estimate for the area between the graph and the x-axis, as illustrated in Figure 6(a). The upper Riemann sum is the sum of the areas of the rectangles with height M_i and width δx_i, giving an *upper* estimate for the area between the graph and the x-axis, as illustrated in Figure 6(b).

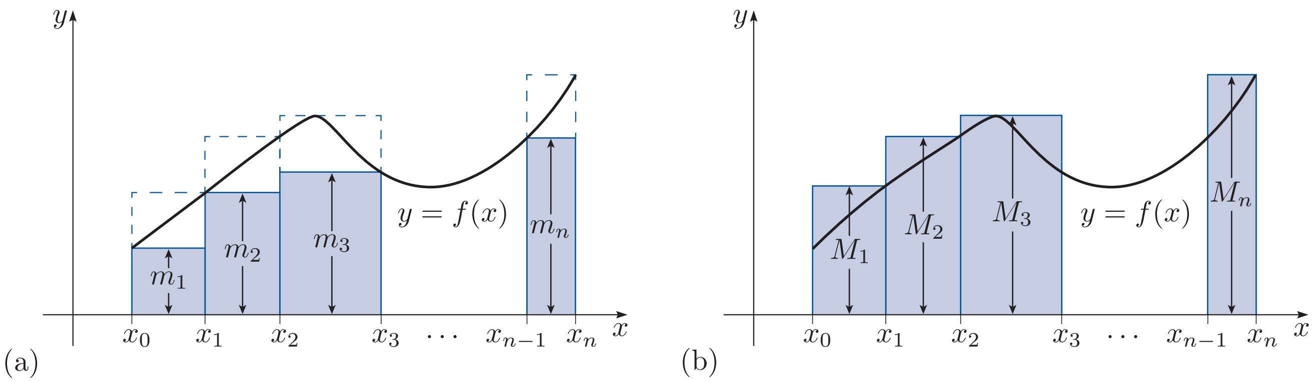


Figure 6 The rectangles whose areas are included in (a) the lower Riemann sum $L(f, P)$ and (b) the upper Riemann sum $U(f, P)$

Now, on any interval, the greatest lower bound of a function f is necessarily less than or equal to its least upper bound. It follows that, in each subinterval $[x_{i-1}, x_i]$, we have $m_i \leq M_i$. Summing from $i = 1$ to n, we obtain the following result.

Theorem F43

Let f be a bounded function on $[a, b]$, and let P be a partition of $[a, b]$. Then

$$L(f, P) \leq U(f, P).$$

Worked Exercise F26

Let

$$f(x) = \begin{cases} 2x, & 0 < x < 1, \\ 1, & x = 0, 1, \end{cases}$$

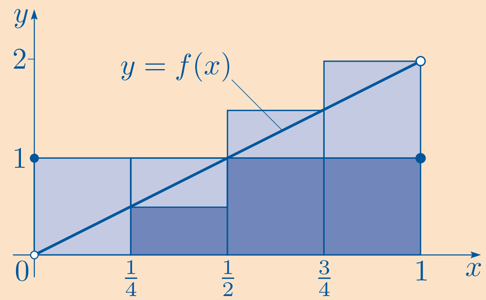
and let

$$P = \left\{ \left[0, \tfrac{1}{4}\right], \left[\tfrac{1}{4}, \tfrac{1}{2}\right], \left[\tfrac{1}{2}, \tfrac{3}{4}\right], \left[\tfrac{3}{4}, 1\right] \right\}$$

be a partition of $[0, 1]$. Determine $L(f, P)$ and $U(f, P)$.

Solution

It is helpful to make a sketch of the graph of f and the partition of the interval $[0, 1]$.



Using the notation in the definitions, we set out the information needed to calculate the lower and upper Riemann sums using a layout that will make the calculation straightforward. Note that this function is increasing except at $x = 0$ and at $x = 1$, so we have to take special care at these points: the values of M_1 and m_4 are not what you might expect, and the values of m_1 and M_4 are not taken by the function.

For the four subintervals in P, we have

$$\begin{aligned}
&m_1 = 0, && M_1 = f(0) = 1, && \delta x_1 = \tfrac{1}{4},\\
&m_2 = f\left(\tfrac{1}{4}\right) = \tfrac{1}{2}, && M_2 = f\left(\tfrac{1}{2}\right) = 1, && \delta x_2 = \tfrac{1}{4},\\
&m_3 = f\left(\tfrac{1}{2}\right) = 1, && M_3 = f\left(\tfrac{3}{4}\right) = \tfrac{3}{2}, && \delta x_3 = \tfrac{1}{4},\\
&m_4 = f(1) = 1, && M_4 = 2, && \delta x_4 = \tfrac{1}{4}.
\end{aligned}$$

Then

$$\begin{aligned}
L(f, P) &= \sum_{i=1}^{4} m_i \delta x_i\\
&= \left(0 \times \tfrac{1}{4}\right) + \left(\tfrac{1}{2} \times \tfrac{1}{4}\right) + \left(1 \times \tfrac{1}{4}\right) + \left(1 \times \tfrac{1}{4}\right)\\
&= 0 + \tfrac{1}{8} + \tfrac{1}{4} + \tfrac{1}{4}\\
&= \tfrac{5}{8}
\end{aligned}$$

and

$$\begin{aligned}
U(f, P) &= \sum_{i=1}^{4} M_i \delta x_i\\
&= \left(1 \times \tfrac{1}{4}\right) + \left(1 \times \tfrac{1}{4}\right) + \left(\tfrac{3}{2} \times \tfrac{1}{4}\right) + \left(2 \times \tfrac{1}{4}\right)\\
&= \tfrac{1}{4} + \tfrac{1}{4} + \tfrac{3}{8} + \tfrac{1}{2}\\
&= \tfrac{11}{8}.
\end{aligned}$$

Exercise F35

Let

$$f(x) = \begin{cases} 2x, & 0 < x < 1, \\ 1, & x = 0, 1, \end{cases}$$

and let

$$P = \{[0, \tfrac{1}{5}], [\tfrac{1}{5}, \tfrac{1}{2}], [\tfrac{1}{2}, 1]\}$$

be a partition of $[0, 1]$. Determine $L(f, P)$ and $U(f, P)$.

Worked Exercise F26 and Exercise F35 addressed the same function f on the same interval $[0, 1]$ but with different partitions. The two lower sums were $5/8$ and $31/50$, and the two upper sums were $11/8$ and $3/2$; each of the two lower sums is smaller than *both* of the two upper sums.

We now look at another example to see whether this happens again in a different situation. We make use of the following formula for a sum of squares which you may have met in your previous studies (it can be proved by mathematical induction):

$$1^2 + 2^2 + \cdots + n^2 = \sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}, \quad \text{for } n = 1, 2, \ldots.$$

Worked Exercise F27

Let

$$f(x) = \begin{cases} x^2, & 0 \le x \le 1, \\ 2, & 1 < x \le 2, \end{cases}$$

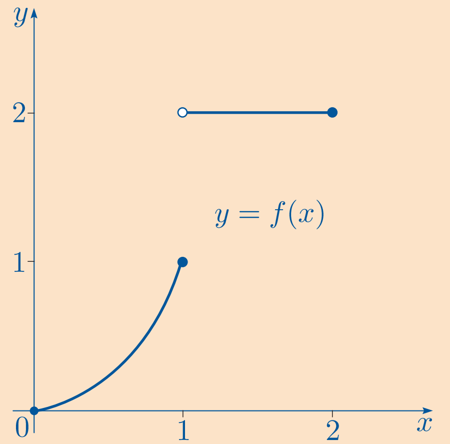
and, for each $n \in \mathbb{N}$, let

$$P_{2n} = \{[0, \tfrac{1}{n}], [\tfrac{1}{n}, \tfrac{2}{n}], \ldots, [2 - \tfrac{1}{n}, 2]\}$$

be the standard partition of $[0, 2]$ into $2n$ equal subintervals. Determine $L(f, P_{2n})$ and $U(f, P_{2n})$.

Solution

The graph of f is shown below.



The function f is increasing on $[0, 2]$. Thus, on each subinterval in $[0, 2]$, the infimum of f is the value of f at the left endpoint of the subinterval and the supremum of f is the value of f at the right endpoint of the subinterval. Since the ith subinterval in P_{2n} is $[x_{i-1}, x_i] = \left[\frac{i-1}{n}, \frac{i}{n}\right]$, we have, for $i = 1, 2, \ldots, 2n$,

$$m_i = f\left(\frac{i-1}{n}\right), \quad M_i = f\left(\frac{i}{n}\right), \quad \delta x_i = \frac{i}{n} - \frac{i-1}{n} = \frac{1}{n}.$$

Now $f(x) = x^2$ on $[0, 1]$ and $f(x) = 2$ on $(1, 2]$ so, since $x_n = 1$, we have

$$m_i = \left(\frac{i-1}{n}\right)^2, \quad \text{for } i = 1, 2, \ldots, n+1,$$

and

$$m_i = 2, \quad \text{for } i = n+2, n+3, \ldots, 2n.$$

Hence

$$\begin{aligned}
L(f, P_{2n}) &= \sum_{i=1}^{2n} m_i \delta x_i = \sum_{i=1}^{n+1} m_i \delta x_i + \sum_{i=n+2}^{2n} m_i \delta x_i \\
&= \left(0 + \left(\frac{1}{n}\right)^2 + \cdots + \left(\frac{n}{n}\right)^2\right) \times \frac{1}{n} + ((n-1) \times 2) \times \frac{1}{n} \\
&= \left(\frac{1^2 + \cdots + n^2}{n^2} \times \frac{1}{n}\right) + 2\frac{n-1}{n} \\
&= \frac{n(n+1)(2n+1)}{6n^3} + 2 - \frac{2}{n}.
\end{aligned}$$

Also, we have

$$M_i = \left(\frac{i}{n}\right)^2, \quad \text{for } i = 1, 2, \ldots, n,$$

and

$$M_i = 2, \quad \text{for } i = n+1, n+2, \ldots, 2n.$$

Hence

$$\begin{aligned}
U(f, P_{2n}) &= \sum_{i=1}^{2n} M_i \delta x_i \\
&= \sum_{i=1}^{n} M_i \delta x_i + \sum_{i=n+1}^{2n} M_i \delta x_i \\
&= \left(\left(\frac{1}{n}\right)^2 + \left(\frac{2}{n}\right)^2 + \ldots + \left(\frac{n}{n}\right)^2\right) \times \frac{1}{n} + (n \times 2) \times \frac{1}{n} \\
&= \left(\frac{1^2 + \ldots + n^2}{n^2} \times \frac{1}{n}\right) + 2 \\
&= \frac{n(n+1)(2n+1)}{6n^3} + 2.
\end{aligned}$$

We now look at the result of Worked Exercise F27 for different values of n. As n increases, the number, $2n$, of subintervals in the partition increases and the length of each subinterval, $1/n$, decreases. From the above formulas for $L(f, P_{2n})$ and $U(f, P_{2n})$, we find that, to three decimal places, the lower and upper Riemann sums are then as given in the following table.

		$L(f, P_{2n})$	$U(f, P_{2n})$
$n = 2$:	4 equal subintervals	1.625	2.625
$n = 4$:	8 equal subintervals	1.969	2.469
$n = 10$:	20 equal subintervals	2.185	2.385
$n = 100$:	200 equal subintervals	2.318	2.338

As the subintervals increase in number and decrease in length, the lower sums increase and the upper sums decrease. But the lower sums are all less than or equal to all the upper sums! In fact this is always the case, as stated in the following result.

Theorem F44

Let f be a bounded function on $[a, b]$, and let P and P' be partitions of $[a, b]$. Then

$$L(f, P) \leq U(f, P').$$

You saw in Theorem F43 that, for a given partition P,

$$L(f, P) \leq U(f, P).$$

In order to prove the general case in Theorem F44 with two different partitions P and P' we require some new ideas. We develop these ideas and prove this result in Subsection 1.4.

The integral

We now return to our original problem: how to define an 'integral' that pins down our intuitive notion of 'the area under a curve'. We have seen that lower Riemann sums provide underestimates for this 'area' and upper Riemann sums provide overestimates; see Figure 7.

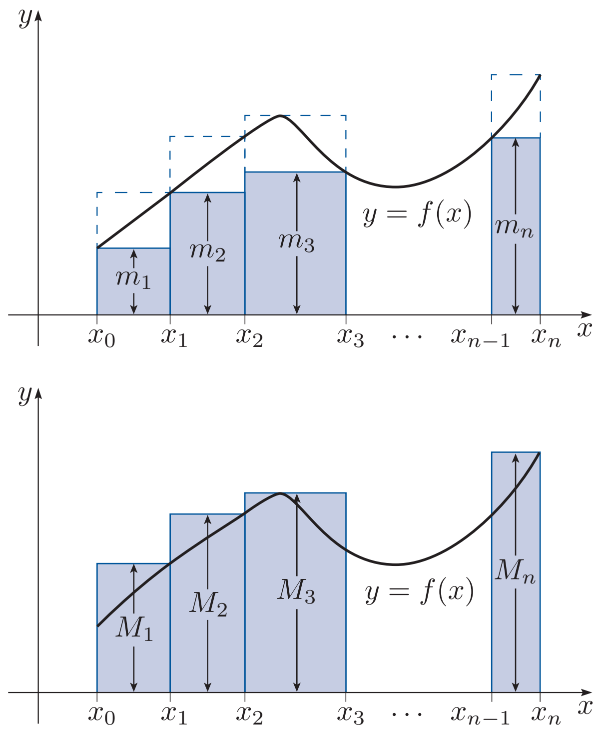


Figure 7 Lower and upper Riemann sums

So we make the following definitions.

Definitions

Let f be a bounded function on a closed interval $[a,b]$, let P be a partition of $[a,b]$ and let $L(f,P)$ and $U(f,P)$, respectively, be the corresponding lower and upper Riemann sums.

Then the **lower integral** of f on $[a,b]$ is

$$\underline{\int_a^b} f = \sup_P L(f,P),$$

and the **upper integral** of f on $[a,b]$ is

$$\overline{\int_a^b} f = \inf_P U(f,P).$$

We say that f is **integrable** on $[a,b]$ if

$$\underline{\int_a^b} f = \overline{\int_a^b} f,$$

and their common value is then called the **integral** of f on $[a,b]$.

The integral is written as $\int_a^b f$ or $\int_a^b f(x)\,dx$, and a and b are called the **limits of integration**.

Remarks

1. The common value of the lower and upper integrals of f (when it exists) is sometimes known as the *Riemann* integral of f, rather than simply as the integral of f.
2. In the definitions, $\sup_P L(f,P)$ is the supremum of the lower Riemann sums over *all possible* partitions P of the interval $[a,b]$. Similarly, $\inf_P U(f,P)$ is the infimum of the upper Riemann sums over all possible partitions P.
3. For any bounded function f on an interval $[a,b]$, it follows from the fact that all lower Riemann sums are less than or equal to all upper Riemann sums (Theorem F44), and from the above definitions, that the lower integral $\underline{\int_a^b} f$ and the upper integral $\overline{\int_a^b} f$ both exist, and that we always have

$$\underline{\int_a^b} f \leq \overline{\int_a^b} f.$$

(We omit the proofs of these facts.) Note, however, that the lower and upper integrals take different values unless f is integrable.

We now return to the function f that we considered earlier in Worked Exercise F27.

Worked Exercise F28

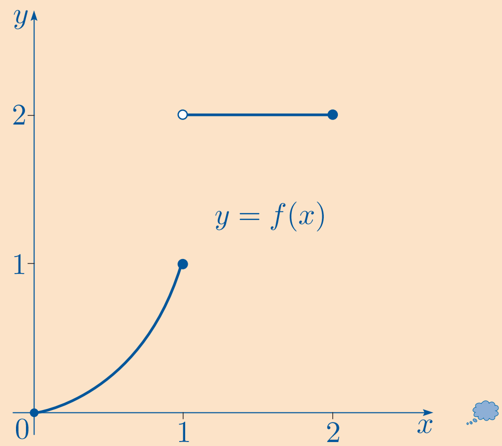
Let

$$f(x) = \begin{cases} x^2, & 0 \leq x \leq 1, \\ 2, & 1 < x \leq 2. \end{cases}$$

Prove that f is integrable on $[0, 2]$, and evaluate $\displaystyle\int_0^2 f$.

Solution

It is helpful to sketch the graph.



We have already seen in Worked Exercise F27 that if we take P_{2n} to be the partition

$$\left\{ \left[0, \frac{1}{n}\right], \left[\frac{1}{n}, \frac{2}{n}\right], \ldots, \left[2 - \frac{1}{n}, 2\right] \right\}$$

of $[0, 2]$, then

$$L(f, P_{2n}) = \frac{n(n+1)(2n+1)}{6n^3} + 2 - \frac{2}{n}$$

and

$$U(f, P_{2n}) = \frac{n(n+1)(2n+1)}{6n^3} + 2.$$

Then, as $n \to \infty$, we have

$$L(f, P_{2n}) = \frac{(1 + 1/n)\,(2 + 1/n)}{6} + 2 - \frac{2}{n} \to \frac{1}{3} + 2 = \frac{7}{3}$$

so that, in particular,

$$\underline{\int_0^2} f \geq \frac{7}{3}.$$

This holds because $\underline{\int_0^2} f$ is defined to be the supremum of the lower Riemann sums of f over all possible partitions of $[0,2]$, so it must be greater than or equal to the particular lower Riemann sum $L(f, P_{2n})$ for any value of n.

Similarly, as $n \to \infty$, we have

$$U(f, P_{2n}) = \frac{(1+1/n)\,(2+1/n)}{6} + 2 \to \frac{1}{3} + 2 = \frac{7}{3}$$

so that

$$\overline{\int_0^2} f \leq \frac{7}{3}.$$

We have now shown that

$$\frac{7}{3} \leq \underline{\int_0^2} f \leq \overline{\int_0^2} f \leq \frac{7}{3}.$$

It follows that

$$\underline{\int_0^2} f = \overline{\int_0^2} f,$$

so f is integrable on $[0,2]$ and $\displaystyle\int_0^2 f = \frac{7}{3}$.

However, not all bounded functions defined on closed intervals are integrable!

Worked Exercise F29

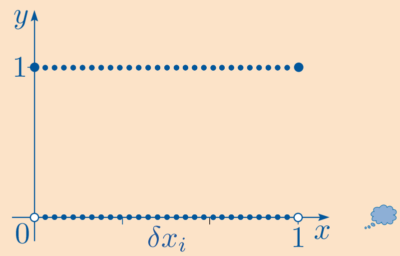
Let f be the Dirichlet function on $[0,1]$ defined by

$$f(x) = \begin{cases} 1, & 0 \leq x \leq 1,\ x \text{ rational} \\ 0, & 0 \leq x \leq 1,\ x \text{ irrational.} \end{cases}$$

Determine the values of $\underline{\int_0^1} f$ and $\overline{\int_0^1} f$, and hence show that f is not integrable on $[0,1]$.

Solution

The graph of f is shown below.



Let $P = \{[x_0, x_1], [x_1, x_2], \ldots, [x_{n-1}, x_n]\}$, where $x_0 = 0$, $x_1 = 1$, be any partition of $[0, 1]$. Then, on each subinterval $[x_{i-1}, x_i]$ in P, we have

$$m_i = 0 \text{ and } M_i = 1, \quad \text{for } i = 1, 2, \ldots, n.$$

This is because every subinterval contains both rational and irrational points, by the density property of the real numbers (see Subsection 1.4 of Unit D1).

So

$$L(f, P) = \sum_{i=1}^{n} m_i \delta x_i = \sum_{i=1}^{n} (0 \times \delta x_i) = 0$$

and

$$U(f, P) = \sum_{i=1}^{n} M_i \delta x_i = \sum_{i=1}^{n} (1 \times \delta x_i) = \sum_{i=1}^{n} \delta x_i = 1,$$

since the sum of the lengths of all the subintervals is equal to the length of the interval $[0, 1]$.

It follows that

$$\underline{\int_0^1} f = \sup_P L(f, P) = 0$$

and

$$\overline{\int_0^1} f = \inf_P U(f, P) = 1.$$

Then, since $\underline{\int_0^1} f \neq \overline{\int_0^1} f$, we conclude that f is not integrable on $[0, 1]$.

Exercise F36

For the function

$$f(x) = x, \quad x \in [0,1],$$

and the standard partition of $[0,1]$

$$P_n = \left\{ \left[0, \frac{1}{n}\right], \left[\frac{1}{n}, \frac{2}{n}\right], \dots, \left[\frac{n-1}{n}, 1\right] \right\}, \quad n \in \mathbb{N},$$

determine the values of $L(f, P_n)$ and $U(f, P_n)$. Hence show whether f is integrable on $[0,1]$ and, if it is, determine the value of $\displaystyle\int_0^1 f$.

1.2 Criteria for integrability

It would be tedious to have to go back to the definition of integrability whenever we wish to show that a given function is integrable on a closed interval. You will now meet a number of criteria that we can use to avoid this.

In order to prove that a bounded function f is integrable on $[a, b]$ directly from the definition of the integral, we need to look at $\sup\limits_P L(f, P)$ and $\inf\limits_P U(f, P)$ over *all* partitions P of $[a, b]$. However, as you saw in Worked Exercise F28 and Exercise F36, in many situations it is sufficient to consider just one *sequence* of partitions (P_n) in order to establish integrability. The conditions under which this simplification holds are set out in the following result. (Recall that $\|P_n\|$, the *mesh* of P_n, is the length of the longest subinterval of P_n.)

Theorem F45

Let f be a bounded function on $[a, b]$. If there is a sequence of partitions (P_n) of $[a, b]$ such that $\|P_n\| \to 0$ and

$$\lim_{n\to\infty} L(f, P_n) = \lim_{n\to\infty} U(f, P_n) = A, \quad \text{where } A \in \mathbb{R},$$

then f is integrable on $[a, b]$ and $\displaystyle\int_a^b f = A$.

Proof Let $\varepsilon > 0$. It follows from the equations in the statement of the theorem that there exists an integer n such that

$$L(f, P_n) > A - \tfrac{1}{2}\varepsilon \quad \text{and} \quad U(f, P_n) < A + \tfrac{1}{2}\varepsilon. \tag{1}$$

Now, by the definitions of upper and lower integrals,

$$\underline{\int_a^b} f \geq L(f, P_n) \quad \text{and} \quad \overline{\int_a^b} f \leq U(f, P_n). \tag{2}$$

Combining inequalities (1) and (2), we obtain

$$A - \tfrac{1}{2}\varepsilon < \underline{\int_a^b} f \leq \overline{\int_a^b} f < A + \tfrac{1}{2}\varepsilon.$$

Here we have used the inequality $\underline{\int_a^b} f \leq \overline{\int_a^b} f$ mentioned in the remarks after the definition of the integral in the previous subsection. As stated there, this follows from Theorem F44.

Since ε is any positive number, we deduce that the upper and lower integrals of f on $[a, b]$ are equal to A, so f is integrable and

$$\int_a^b f = A.$$

■

In fact the following result in the opposite direction to Theorem F45 also holds. However, its proof is somewhat more complicated, and we defer this to Subsection 1.4.

Theorem F46

If f is an integrable function on $[a, b]$ and (P_n) is a sequence of partitions of $[a, b]$ such that $\|P_n\| \to 0$, then

$$\lim_{n\to\infty} L(f, P_n) = \lim_{n\to\infty} U(f, P_n) = \int_a^b f.$$

This result is particularly useful for proving that a function is *not* integrable. For if f is defined on $[a, b]$ and we can find a sequence of partitions (P_n) of $[a, b]$ whose mesh tends to zero but for which

$$\lim_{n\to\infty} L(f, P_n) \neq \lim_{n\to\infty} U(f, P_n),$$

then it follows from Theorem F46 that f is not integrable on $[a, b]$.

Exercise F37

For each of the following functions f, determine whether f is integrable on $[0, 1]$ and, if it is, find $\int_0^1 f$.

(a) $f(x) = \begin{cases} -2, & 0 \leq x < 1, \\ 3, & x = 1. \end{cases}$

(b) $f(x) = \begin{cases} x, & 0 \leq x \leq 1, \quad x \text{ rational}, \\ 0, & 0 \leq x \leq 1, \quad x \text{ irrational}. \end{cases}$

The next result is of particular interest in that its statement says nothing about the value of the integral itself: it mentions only the difference between the lower and the upper Riemann sums. The result follows from Theorems F45 and F46.

Corollary F47 Riemann's Criterion

Let f be bounded on $[a, b]$. Then

f is integrable on $[a, b]$

if and only if

there is a sequence (P_n) of partitions of $[a, b]$ with $\|P_n\| \to 0$ such that $U(f, P_n) - L(f, P_n) \to 0$.

Proof Theorem F46 implies that if f is integrable on $[a, b]$, and (P_n) is a sequence of partitions of $[a, b]$ with $\|P_n\| \to 0$, then

$$U(f, P_n) - L(f, P_n) \to 0.$$

On the other hand, if there is a sequence (P_n) of partitions of $[a, b]$ with $\|P_n\| \to 0$ such that $U(f, P_n) - L(f, P_n) \to 0$, then, because

$$L(f, P_n) \leq \underline{\int_a^b} f \leq \overline{\int_a^b} f \leq U(f, P_n),$$

these upper and lower integrals must be equal, so f is integrable on $[a, b]$, by Theorem F45. ■

Two classes of integrable functions

With Riemann's Criterion at our disposal, we can now determine some large classes of functions that are always integrable: the monotonic functions and the continuous functions.

Theorem F48

A function f which is bounded and monotonic on $[a, b]$ is integrable on $[a, b]$.

Proof We prove this theorem in the case when f is increasing on $[a, b]$. (The proof is similar if f is decreasing.)

Consider the standard partition of $[a, b]$; that is,

$$P_n = \{[x_0, x_1], [x_1, x_2], \ldots, [x_{n-1}, x_n]\}, \tag{3}$$

where

$$x_i = a + i\,\frac{b-a}{n}, \quad \text{for } i = 0, 1, 2, \ldots, n. \tag{4}$$

Now f is increasing, so on each subinterval $[x_{i-1}, x_i]$, for $i = 1, 2, \ldots, n$,

$$m_i = f(x_{i-1}) \quad \text{and} \quad M_i = f(x_i);$$

see Figure 8. Also, $\delta x_i = (b-a)/n$, for $i = 1, 2, \ldots, n$. Hence

$$\begin{aligned} U(f, P_n) - L(f, P_n) &= \sum_{i=1}^{n} (M_i - m_i)\, \delta x_i \\ &= \frac{b-a}{n} \sum_{i=1}^{n} (f(x_i) - f(x_{i-1})) \\ &= \frac{b-a}{n} (f(x_n) - f(x_0)) \\ &= \frac{b-a}{n} (f(b) - f(a)). \end{aligned}$$

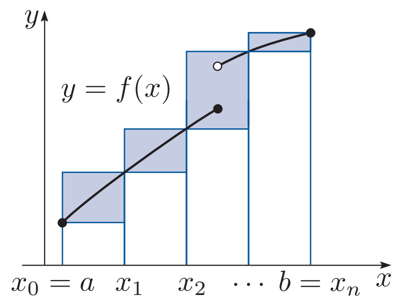


Figure 8 The standard partition for an increasing function f

Here we have used the fact that the terms of the series are

$$f(x_1) - f(x_0), \quad f(x_2) - f(x_1), \quad \ldots, \quad f(x_n) - f(x_{n-1}),$$

so everything except $f(x_0)$ and $f(x_n)$ cancels.

The sequence $((b-a)(f(b) - f(a))/n)$ is null, so it follows from Riemann's Criterion that f is integrable on $[a, b]$. ■

Theorem F49

A function f which is continuous on $[a, b]$ is integrable on $[a, b]$.

Proof We use the fact that f must be uniformly continuous on $[a, b]$.

You saw in Theorem F19 in Subsection 4.2 of Unit F1 that a function which is continuous on a bounded closed interval is uniformly continuous there.

Let $\varepsilon > 0$. Then there exists $\delta > 0$ such that

$$|f(x) - f(y)| < \frac{\varepsilon}{b-a}, \quad \text{for all } x, y \in [a, b] \text{ with } |x - y| < \delta. \tag{5}$$

We use $\varepsilon/(b-a)$ here in order to obtain ε later in the proof.

Next we choose $N \in \mathbb{N}$ such that $(b-a)/N < \delta$. For $n \geq N$, let P_n be the standard partition of $[a, b]$ given by equations (3) and (4).

Now f is continuous on each subinterval $[x_{i-1}, x_i]$, for $i = 1, 2, \ldots, n$. Thus, by the Extreme Value Theorem (see Subsection 3.3 of Unit D4), there exist points c_i and d_i in $[x_{i-1}, x_i]$ such that

$$m_i = f(c_i) \quad \text{and} \quad M_i = f(d_i). \tag{6}$$

Since $[x_{i-1}, x_i]$ has length $(b-a)/n < \delta$, we deduce by statements (5) and (6) that

$$M_i - m_i < \frac{\varepsilon}{b-a}.$$

Here we have used the fact that $|d_i - c_i| < \dfrac{b-a}{n} < \delta$, from which it follows that $|f(d_i) - f(c_i)| = M_i - m_i < \dfrac{\varepsilon}{b-a}$ by statement (5).

Hence, for $n \geq N$ we have

$$\begin{aligned} U(f, P_n) - L(f, P_n) &= \sum_{i=1}^{n} (M_i - m_i)\,\delta x_i \\ &< \sum_{i=1}^{n} \left(\frac{\varepsilon}{b-a}\right)\left(\frac{b-a}{n}\right) = \varepsilon. \end{aligned}$$

Thus $(U(f, P_n) - L(f, P_n))$ is a null sequence, so it follows from Riemann's Criterion that f is integrable on $[a, b]$. ■

Theorems F48 and F49 show that monotonic functions and continuous functions are integrable, but we know that some bounded functions are not integrable: for example, the Dirichlet function, as you saw in Worked Exercise F29. This suggests the question: precisely which bounded functions are integrable?

The full answer to this question is rather complicated but, roughly speaking, a bounded function is integrable on $[a, b]$ if and only if it is continuous at 'most' points of $[a, b]$. However, it is possible for a function to be discontinuous at infinitely many points of $[a, b]$ and yet be integrable on $[a, b]$. For example, the Riemann function which you met in Section 3 of Unit F1 is discontinuous at all rational points and yet it can be shown to be integrable on $[0, 1]$, the value of its integral being 0 (we do not prove this here).

Georg Friedrich Bernhard Riemann

Riemann and Lebesgue integration

Georg Friedrich Bernhard Riemann (1826–1866) laid down the fundamental ideas of the integral that is now named after him when he was writing his doctoral thesis in 1854. This was published posthumously in 1867. His formulation and proof were rather obscure and the version that is generally used today (including in this module) was given by the French mathematician Gaston Darboux (1842–1917) in 1875.

In 1902 a different definition of the integral was given by another French mathematician, Henri Léon Lebesgue (1875–1941). The main difference between the two definitions is in the way the area under the curve is measured. The Riemann integral considers the area as being made up of vertical rectangles, while the Lebesgue integral considers horizontal rectangles. Or to put it another way, the Riemann integral considers the domain of the function while the Lebesgue integral considers the codomain of the function. Although Lebesgue's definition has the advantage that it is applicable to a larger class of functions than Riemann's definition, it requires the formal notion of a measure. (You may study measure theory in the future, but this topic is not covered in M208.)

Lebesgue himself provided a rather nice example to illustrate the difference between his approach and that of Riemann:

> I have to pay a certain sum, which I have collected in my pocket. I take the bills and coins out of my pocket and give them to the creditor in the order I find them until I have reached the total sum. This is the Riemann integral. But I can proceed differently. After I have taken out all my money I order the bills and coins according to identical values and then I pay the several heaps one after another to the creditor. This is my integral.

(Source: Siegmund-Schultze, R. (2008) 'Henri Lebesgue', in Gower, T. (ed) *The Princeton Companion to Mathematics*, Princeton, Princeton University Press, p. 796.)

Henri Léon Lebesgue

1.3 Properties of integrals

In your previous study of integration (for example, in a calculus course) you will have met many properties of integrals without a clear explanation of exactly why they hold. We now look at several of these properties and in some cases give an outline of how they follow from the definition of an integral in Subsection 1.1 and the various theorems that you met in Subsection 1.2. In reading this subsection, you should concentrate on understanding the various properties themselves; our comments on why the properties hold are optional reading in case you are interested.

First, we look at the limits of the integral. We have already defined integrals of the form $\int_a^b f$, where $a < b$; we now look at the situation where $a = b$ or $a > b$.

Definitions

Let f be a bounded function that is integrable on an interval I containing a and b, where $a < b$. Then we make the following definitions.

- $\int_a^a f = 0$
- $\int_b^a f = -\int_a^b f$

Thus, for example, $\int_1^0 x\,dx$ is defined to equal $-\int_0^1 x\,dx$; you have already seen in Exercise F36 that $\int_0^1 x\,dx = \frac{1}{2}$, so we define the value of $\int_1^0 x\,dx$ to be $-\frac{1}{2}$.

Next we look at the integrability of a bounded function on intervals with endpoints a, b and c, irrespective of the order of these endpoints on the x-axis. Figure 9 illustrates two possibilities: $a < c < b$ and $a < b < c$. The following result applies whatever the order of a, b and c.

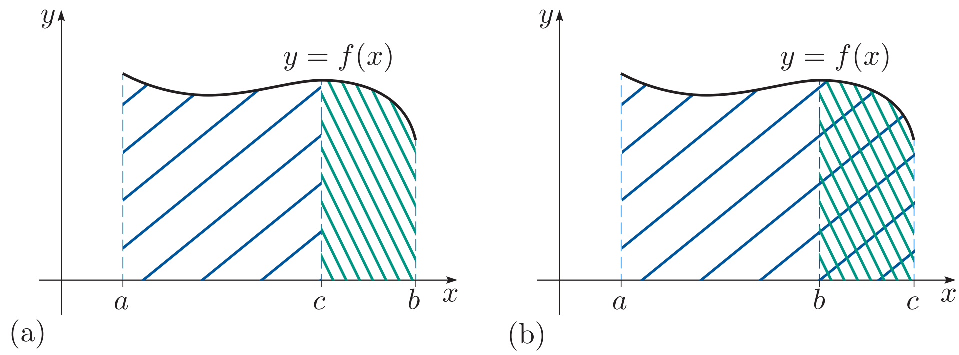


Figure 9 Integrating a function on intervals with endpoints a, b and c (a) when $a < c < b$ and (b) when $a < b < c$

Theorem F50 Additivity of integrals

Let f be a bounded function that is integrable on an interval I containing the points a, b and c. Then

$$\int_a^c f + \int_c^b f = \int_a^b f.$$

This result can be proved directly from the definitions above, and the definition of the integral given in Subsection 1.1. Notice that, in the situation illustrated in Figure 9(b), $\int_c^b f$ is negative since $b < c$.

Our next result says that the integral of a non-negative function is non-negative, and the integral of a non-positive function is non-positive; see Figure 10. This result can also be proved directly from the definition of the integral.

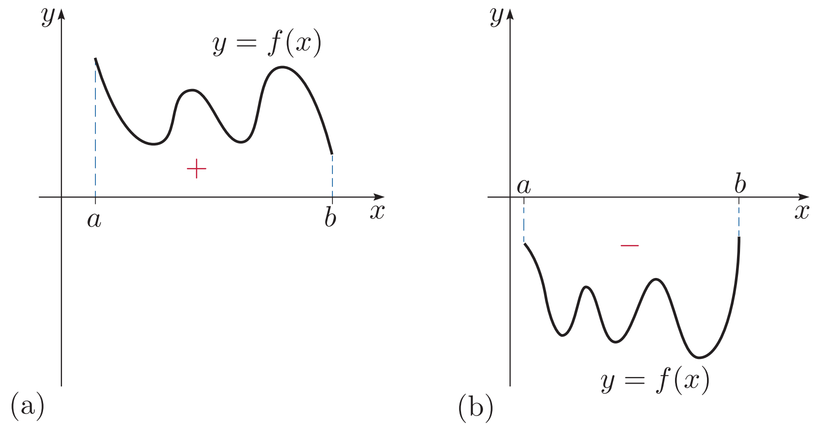


Figure 10 The sign of the integral for (a) a non-negative function and (b) a non-positive function

Theorem F51 Sign of an integral

Let f be a bounded function that is integrable on $[a, b]$.

- If $f(x) \geq 0$ on $[a, b]$, then $\displaystyle\int_a^b f \geq 0$.
- If $f(x) \leq 0$ on $[a, b]$, then $\displaystyle\int_a^b f \leq 0$.

Of course, a function can be non-negative on some parts of its domain and non-positive on other parts. In these circumstances, subintervals where the function takes negative values make a negative contribution to the total area between the graph of the function and the x-axis. If the interval $[a, b]$ on which a function is defined is made up of finitely many subintervals where the function is either always non-negative or always non-positive, then we can evaluate the integral $\displaystyle\int_a^b f$ by applying Theorem F50. If there are infinitely many such subintervals, then the evaluation of the integral involves summing an infinite series; we do not pursue this here.

Next we note that if a function f is integrable, then so is $|f|$; see Figure 11.

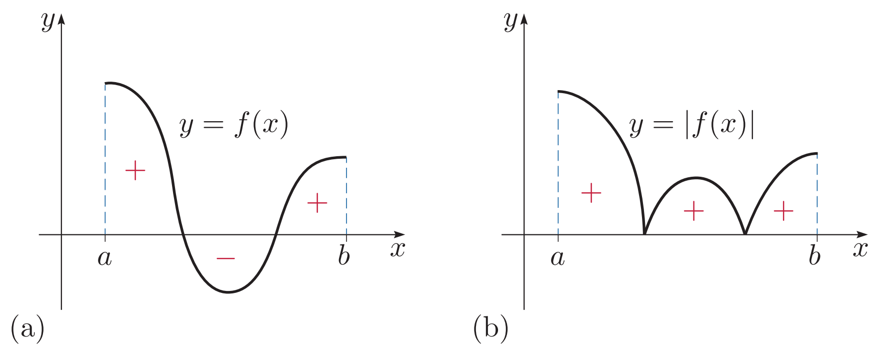


Figure 11 (a) The integral of f (b) The integral of $|f|$

Theorem F52 Modulus Rule

If f is integrable on $[a, b]$, then $|f|$ is also integrable on $[a, b]$.

Outline of the proof (optional)

For a bounded function f on $[a, b]$ and a partition of $[a, b]$

$$P = \{[x_0, x_1], [x_1, x_2], \dots, [x_{n-1}, x_n]\},$$

we define, for $i = 1, 2, \dots, n$, the **variation** $\omega_i(f)$ of f over the subinterval $[x_{i-1}, x_i]$ to be

$$\omega_i(f) = \sup\{|f(x) - f(y)| : x, y \in [x_{i-1}, x_i]\}.$$

It can be shown that

$$\omega_i(f) = \sup\{f(x) : x \in [x_{i-1}, x_i]\} - \inf\{f(x) : x \in [x_{i-1}, x_i]\}.$$

(We omit the details here.)

Hence

$$U(f, P) - L(f, P) = \sum_{i=1}^{n} \omega_i(f)\delta x_i, \tag{7}$$

where $\delta x_i = x_i - x_{i-1}$, as usual.

Now, by the backwards form of the Triangle Inequality which you met in Subsection 3.1 of Unit D1, we have

$$||f(x)| - |f(y)|| \le |f(x) - f(y)|, \quad \text{for } x, y \in [x_{i-1}, x_i],$$

so that

$$\omega_i(|f|) \le \omega_i(f), \quad \text{for } i = 1, 2, \dots, n.$$

Hence, by equation (7),

$$U(|f|, P) - L(|f|, P) \le U(f, P) - L(f, P), \tag{8}$$

for any partition P of $[a, b]$. We can now use Riemann's Criterion to deduce from inequality (8) that if f is integrable on $[a, b]$, then so is $|f|$. ■

Finally, we set out the Combination Rules for integrable functions; we use these a great deal to construct 'new integrable functions from old'.

Theorem F53 Combination Rules for integrable functions

If f and g are integrable on $[a, b]$, then so are the following functions.

Sum Rule $f + g$, with integral

$$\int_a^b (f+g) = \int_a^b f + \int_a^b g$$

Multiple Rule λf, for $\lambda \in \mathbb{R}$, with integral

$$\int_a^b \lambda f = \lambda \int_a^b f$$

Product Rule fg

Quotient Rule f/g, provided that $1/g$ is bounded on $[a, b]$.

You will meet some techniques for finding the integrals of products and quotients in the next section.

The Combination Rules can be proved using the same approach as for the Modulus Rule. The proof uses the following inequalities, which relate the variations of the new functions over a subinterval $[x_{i-1}, x_i]$ of a partition to those of the known integrable functions f and g:

- $\omega_i(f + g) \leq \omega_i(f) + w_i(g)$
- $\omega_i(\lambda f) \leq |\lambda|\omega_i(f),\ \text{for } \lambda \in \mathbb{R}$
- $\omega_i(fg) \leq M(\omega_i(f) + \omega_i(g)),\ \text{where } M = \max\{\sup|f|, \sup|g|\}.$

1.4 Proofs of Theorem F44 and Theorem F46 (optional)

In the proofs of these results, we will use the notion of a *refinement* of a partition. If P is a partition of an interval $[a, b]$, then any partition obtained from P by adding to it a finite number of partition points is called a **refinement** of P. The partition of $[a, b]$ obtained from two partitions P and P' of $[a, b]$ by using all their partition points is called the **common refinement** of P and P'.

For example, for the partitions

$$P = \{[0, \tfrac{1}{4}], [\tfrac{1}{4}, \tfrac{1}{2}], [\tfrac{1}{2}, \tfrac{3}{4}], [\tfrac{3}{4}, 1]\}$$

and

$$P' = \{[0, \tfrac{1}{3}], [\tfrac{1}{3}, \tfrac{2}{3}], [\tfrac{2}{3}, 1]\}$$

of $[0, 1]$, the common refinement of P and P' is the partition

$$\{[0, \tfrac{1}{4}], [\tfrac{1}{4}, \tfrac{1}{3}], [\tfrac{1}{3}, \tfrac{1}{2}], [\tfrac{1}{2}, \tfrac{2}{3}], [\tfrac{2}{3}, \tfrac{3}{4}], [\tfrac{3}{4}, 1]\},$$

as illustrated in Figure 12.

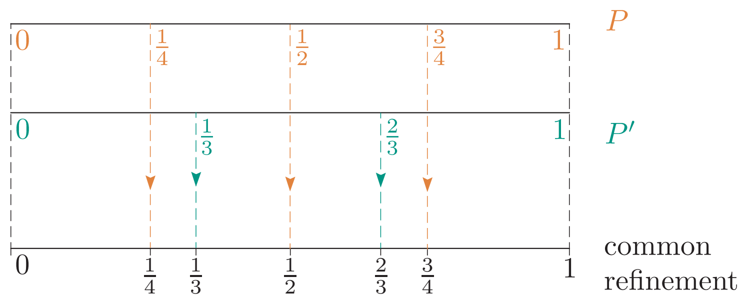


Figure 12 The common refinement of two partitions of $[0, 1]$

We first prove Theorem F44, which is restated below.

Theorem F44

Let f be a bounded function on $[a, b]$, and let P and P' be partitions of $[a, b]$. Then

$$L(f, P) \leq U(f, P').$$

Proof In our proof we will assume that f is non-negative on $[a, b]$. The general result can be deduced by applying the 'non-negative version' of the result to the function $g = f + c$, where c is a constant so large that g is non-negative on $[a, b]$; such a constant c exists since f is bounded on $[a, b]$.

First recall that, from Theorem F43, for any partition P of $[a, b]$ we have

$$L(f, P) \leq U(f, P).$$

We will prove Theorem F44 by showing that if P'' is the common refinement of P and P', then

$$L(f, P) \leq L(f, P'') \leq U(f, P'') \leq U(f, P'). \tag{9}$$

We claim that adding a new partition point x' to a partition

$$P = \{[x_0, x_1], [x_1, x_2], \ldots, [x_{n-1}, x_n]\}$$

of $[a, b]$ does not increase the upper Riemann sum and may decrease it. This is because the new point x' lies in a subinterval $[x_{i-1}, x_i]$, for some i, and the only effect on the upper Riemann sum of adding x' is to replace the rectangle with side M_i standing on $[x_{i-1}, x_i]$ with a pair of adjacent rectangles standing on $[x_{i-1}, x_i]$ with heights at most M_i, as shown in Figure 13.

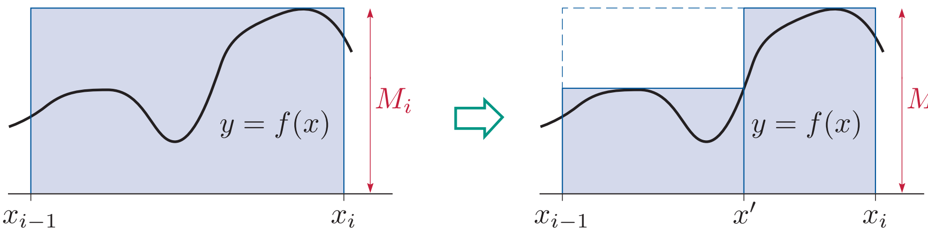


Figure 13 The effect on the upper Riemann sum of adding a new partition point

Hence adding a finite number of new partition points to P does not increase the upper Riemann sum. Similarly, adding a finite number of new partition points to P does not decrease the lower Riemann sum.

Now the partition P'' can be formed from either P or P' by adding a finite number of partition points, so inequalities (9) follow. ■

To end this subsection we give the proof of Theorem F46.

Theorem F46

If f is an integrable function on $[a, b]$ and (P_n) is a sequence of partitions of $[a, b]$ such that $\|P_n\| \to 0$, then

$$\lim_{n\to\infty} L(f, P_n) = \lim_{n\to\infty} U(f, P_n) = \int_a^b f.$$

Proof We prove that $\lim_{n\to\infty} U(f, P_n) = \int_a^b f$. The proof that $\lim_{n\to\infty} L(f, P_n) = \int_a^b f$ is similar.

We assume that f is non-negative on $[a, b]$ so, for some $M \in \mathbb{R}$, we have

$$0 \le f(x) \le M, \quad \text{for } a \le x \le b. \tag{10}$$

Let $\varepsilon > 0$. Since f is integrable on $[a, b]$, there is a partition

$$P' = \{[x'_0, x'_1], [x'_1, x'_2], \ldots, [x'_{m-1}, x'_m]\}$$

of $[a, b]$ where $x'_0 = a$ and $x'_m = b$, with m subintervals, such that

$$U(f, P') < \int_a^b f + \tfrac{1}{2}\varepsilon. \tag{11}$$

We use $\frac{1}{2}\varepsilon$ here in order to obtain ε later in the proof.

Now consider any partition in the sequence (P_n), of the form

$$P_n = \{[x_0, x_1], [x_1, x_2], \ldots, [x_{p-1}, x_p]\},$$

where p is the number of subintervals in P_n. For $k = 1, 2, \ldots, p$, we define

$$M_k = \sup\{f(x) : x_{k-1} \le x \le x_k\} \quad \text{and} \quad \delta x_k = x_k - x_{k-1}.$$

Let P'_n denote the common refinement of P' and P_n. Then, as in the proof of Theorem F44, we have

$$U(f, P'_n) \le U(f, P'). \tag{12}$$

Now we can obtain P_n from P'_n by removing at most $m - 1$ of the partition points of P'.

This is because P' has $m + 1$ partition points, and all partitions have the points a and b in common.

Removing such a point x'_i, lying in (x_{k-1}, x_k) say, we increase the upper Riemann sum by at most $M_k(x_k - x_{k-1})$, as illustrated in Figure 14. So, since $M_k \le M$, by inequality (10), and $x_k - x_{k-1} \le \|P_n\|$, for $k = 1, 2, \ldots, p$,

$$U(f, P_n) \le U(f, P'_n) + (m - 1)M\|P_n\|. \tag{13}$$

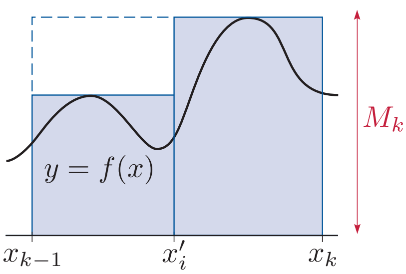


Figure 14 Removing a point x'_i of a partition

Combining inequalities (11), (12) and (13), we obtain

$$U(f, P_n) < \left(\int_a^b f + \tfrac{1}{2}\varepsilon\right) + (m - 1)M\|P_n\|.$$

Since $\|P_n\| \to 0$ as $n \to \infty$, we can choose N so large that

$$(m - 1)M\|P_n\| < \tfrac{1}{2}\varepsilon, \quad \text{for all } n > N.$$

We use $\frac{1}{2}\varepsilon$ here in order to obtain ε in the next step of the proof.

Hence

$$U(f, P_n) < \left(\int_a^b f + \tfrac{1}{2}\varepsilon\right) + \tfrac{1}{2}\varepsilon = \int_a^b f + \varepsilon, \quad \text{for all } n > N.$$

Since $U(f, P_n) \ge \displaystyle\int_a^b f$, by the definition of the integral, we deduce that

$$\left|U(f, P_n) - \int_a^b f\right| = U(f, P_n) - \int_a^b f < \varepsilon, \quad \text{for all } n > N.$$

So $U(f, P_n) - \displaystyle\int_a^b f$ tends to zero as n tends to ∞.

Hence $\displaystyle\lim_{n\to\infty} U(f, P_n) = \int_a^b f$, as required. ■

2 Evaluation of integrals

In Section 1 you saw what it means for a bounded function defined on a closed interval to be integrable. In this section you will study a variety of techniques for evaluating the integrals of such functions.

2.1 The Fundamental Theorem of Calculus

In this subsection we show that integration and differentiation are intimately related by proving a result known as the Fundamental Theorem of Calculus. You probably know this result from a previous course on calculus, but it is worth pausing to reflect on how remarkable it is. In Section 1 we defined the integral by means of more and more accurate estimates of the area between the graph of a function and the x-axis. At no point did it seem that this process was related to differentiation, but the Fundamental Theorem of Calculus shows that integration and differentiation are in some sense inverse processes. As you will see, this fact is enormously helpful in evaluating integrals.

We begin our exploration of these ideas by defining a *primitive* of a function.

Definition

Let f be a function defined on an interval I. Then a function F is a **primitive** of f on I if F is differentiable on I and

$$F'(x) = f(x), \quad \text{for } x \in I.$$

It follows from this definition that finding a primitive is the inverse of finding a derivative. (For this reason, a primitive is sometimes called an *antiderivative*.) Note that the domain of the primitive F may be larger than the interval I on which f is defined.

As an example of finding a primitive, let

$$f(x) = \tan x.$$

Then the function

$$F(x) = \log(\sec x)$$

is a primitive of f on the interval $(-\pi/2, \pi/2)$, since

$$F'(x) = \frac{1}{\sec x} \sec x \tan x = \tan x, \quad \text{for } x \in (-\pi/2, \pi/2).$$

Exercise F38

(a) Let

$$f(x) = (x^2 - 4)^{-1/2} \quad (x \in (2, \infty)).$$

Prove that

$$F(x) = \log\left(x + (x^2 - 4)^{1/2}\right)$$

is a primitive of f on $(2, \infty)$.

(b) Let

$$f(x) = \operatorname{sech} x \left(= \frac{1}{\cosh x}\right).$$

Prove that

$$F(x) = \tan^{-1}(\sinh x)$$

is a primitive of f on $\mathbb{R}$.

We now state and prove our main result, the Fundamental Theorem of Calculus. The theorem tells us that we can evaluate the integral of a function f on an interval $[a, b]$ by finding a primitive F of f on $[a, b]$. Note that the expression $F(b) - F(a)$ is sometimes written as $\big[F(x)\big]_a^b$ or $F(x)\big|_a^b$.

Theorem F54 Fundamental Theorem of Calculus

Let f be integrable on $[a, b]$ and let F be a primitive of f on $[a, b]$. Then

$$\int_a^b f = F(b) - F(a).$$

Proof Let

$$P_n = \{[x_0, x_1], \dots, [x_{i-1}, x_i], \dots, [x_{n-1}, x_n]\}, \quad \text{for } n = 1, 2, \dots,$$

be a sequence of partitions of $[a, b]$, with $x_0 = a$, $x_n = b$ and $\|P_n\| \to 0$.

On each subinterval

$$[x_{i-1}, x_i], \quad \text{for } i = 1, 2, \dots, n,$$

the function F satisfies the conditions of the Mean Value Theorem which you met in Subsection 4.1 of Unit F2 *Differentiation*, since a primitive is differentiable and hence continuous. Thus there exists a point $c_i \in (x_{i-1}, x_i)$ such that

$$\begin{aligned} F(x_i) - F(x_{i-1}) &= F'(c_i)(x_i - x_{i-1}) \\ &= f(c_i)\,\delta x_i, \end{aligned} \tag{14}$$

where $\delta x_i = x_i - x_{i-1}$.

Now recall that we use the notation $m_i = \inf\{f(x) : x \in [x_{i-1}, x_i]\}$ and $M_i = \sup\{f(x) : x \in [x_{i-1}, x_i]\}$.

Since

$$m_i \le f(c_i) \le M_i, \quad \text{for } i = 1, 2, \ldots, n,$$

it follows that

$$\sum_{i=1}^{n} m_i\,\delta x_i \le \sum_{i=1}^{n} f(c_i)\,\delta x_i \le \sum_{i=1}^{n} M_i\,\delta x_i.$$

Using equation (14), we can rewrite this statement as

$$L(f, P_n) \le \sum_{i=1}^{n} (F(x_i) - F(x_{i-1})) \le U(f, P_n).$$

We now use telescopic cancellation to evaluate the series in these inequalities, which is

$$(F(x_1) - F(x_0)) + (F(x_2) - F(x_1)) + \cdots + (F(x_n) - F(x_{n-1})).$$

The only remaining terms after cancellation are $F(x_n)$ and $-F(x_0)$.

The series has sum $F(x_n) - F(x_0) = F(b) - F(a)$, so

$$L(f, P_n) \le F(b) - F(a) \le U(f, P_n). \tag{15}$$

Since f is integrable on $[a, b]$, the sequences $(L(f, P_n))$ and $(U(f, P_n))$ both converge to $\int_a^b f$, by Theorem F46. It follows from inequalities (15) and the Limit Inequality Rule for sequences (Theorem D11 in Unit D2 *Sequences*) that

$$\int_a^b f \le F(b) - F(a) \le \int_a^b f,$$

which gives the required result. ■

Theorem F54 shows the close relationship between the integral $\int_a^b f$ and any primitive F of f on $[a, b]$. Because of this result, a primitive F is also called an **indefinite integral** of f and denoted by $\int f(x)\,dx$. Moreover, the process of finding a primitive of f is often informally called *integrating* f, and in this context the function f is called an **integrand**. Also, the integral $\int_a^b f$ is often referred to as the **definite integral** of f over $[a, b]$.

We can use Theorem F54 and the table of standard primitives at the end of this unit to evaluate many integrals.

Worked Exercise F30

Evaluate $\int_0^1 2^x \, dx$.

Solution

Here the integrand is the function $x \longmapsto 2^x$; see the table at the end of this unit.

The function $f(x) = 2^x$ has the following primitive on $[0, 1]$:

$$F(x) = 2^x / \log 2.$$

Hence, by the Fundamental Theorem of Calculus,

$$\int_0^1 2^x \, dx = \left[\frac{2^x}{\log 2}\right]_0^1 = \frac{2}{\log 2} - \frac{1}{\log 2} = \frac{1}{\log 2}.$$

Exercise F39

Using the Fundamental Theorem of Calculus and the table of standard primitives, evaluate the following integrals.

(a) $\int_0^4 (x^2 + 9)^{1/2} \, dx$ (b) $\int_1^e \log x \, dx$

2.2 Primitives

It is natural to ask: can a function have more than one primitive on an interval? The answer to this question is 'yes': for example, on $(-1, 1)$ the functions

$$x \longmapsto x^2 \quad \text{and} \quad x \longmapsto x^2 + 1$$

are both primitives of the function

$$x \longmapsto 2x.$$

However, any two primitives of a function f on an interval can differ only by a constant.

Theorem F55 Uniqueness Theorem for Primitives

Let F_1 and F_2 be primitives of f on an interval I. Then there exists some constant c such that

$$F_2(x) = F_1(x) + c, \quad \text{for } x \in I.$$

Proof Since F_1 and F_2 are primitives of f on I,

$$F_1'(x) = f(x) \quad \text{and} \quad F_2'(x) = f(x), \quad \text{for } x \in I,$$

so

$$F_2'(x) - F_1'(x) = 0, \quad \text{for } x \in I.$$

Thus, by the Zero Derivative Theorem (Corollary F39 in Unit F2), there exists a constant c such that

$$F_2(x) - F_1(x) = c, \quad \text{for } x \in I.$$

■

The range of primitives we can find is considerably extended by the use of several Combination Rules. These rules can be proved using the corresponding rules for derivatives; we omit the details.

Theorem F56 Combination Rules for primitives

Let F and G be primitives of f and g, respectively, on an interval I, and let $\lambda \in \mathbb{R}$. Then, on I:

Sum Rule $f + g$ has a primitive $F + G$

Multiple Rule λf has a primitive λF

Scaling Rule $x \longmapsto f(\lambda x)$ has a primitive $x \longmapsto \dfrac{1}{\lambda} F(\lambda x)$, for $\lambda \neq 0$.

For example, it follows from the table of standard primitives and the Combination Rules that the function with domain $\mathbb{R}^+$ and rule

$$x \longmapsto 3x^{-1} + \sinh 2x$$

has a primitive

$$x \longmapsto 3 \log x + \tfrac{1}{2} \cosh 2x.$$

In applications of these Combination Rules we do not usually mention the rules explicitly.

Exercise F40

Using the table of standard primitives and the Combination Rules, find a primitive of each of the following functions.

(a) $f(x) = 4 \log x - 2/(4 + x^2) \quad (x \in (0, \infty))$

(b) $f(x) = 2 \tan 3x + e^{2x} \cos x \quad (x \in (-\pi/6, \pi/6))$

2.3 Techniques of integration

The Fundamental Theorem of Calculus provides a powerful method for evaluating certain integrals. However, even when we know that a function f has a primitive F, it may not be possible to determine F explicitly.

In fact, most functions f, even quite simple ones, have primitives which are not standard functions. For example, the primitives

$$\int e^{-x^2}\,dx \quad \text{and} \quad \int \frac{dx}{(\log x)^2}$$

cannot be expressed as a combination of a finite number of rational, nth root, trigonometric, exponential and logarithmic functions. (Note that here we have used a standard shorthand and written $\displaystyle\int \frac{dx}{(\log x)^2}$ instead of $\displaystyle\int \frac{1}{(\log x)^2}\,dx$.)

However, by using the Combination Rules we can certainly integrate any *polynomial* function, and the primitive is then always another polynomial function. For example,

$$\int (x^2 - x + 5)\,dx = \tfrac{1}{3}x^3 - \tfrac{1}{2}x^2 + 5x.$$

There is also is a standard procedure called the method of *partial fractions* for integrating a *rational* function. This procedure is not used in this module but is often used for evaluating the integrals of rational functions in complex analysis. Using this method, it can be shown that any primitive of a rational function can always be expressed in terms of rational functions, logarithms of rational functions and inverse tangents of linear functions.

There are also various techniques which can be applied to certain other types of function; the art of integration lies in recognising these types. We now describe briefly some basic techniques of integration.

Integration by substitution

We describe two related techniques of integration by substitution. The first is used when the integrand is of the form

$$x \longmapsto f(g(x))g'(x).$$

In this case, if F is a primitive of f, then

$$\frac{d}{dx}F(g(x)) = F'(g(x))g'(x) = f(g(x))g'(x),$$

by the Composition Rule for derivatives (also called the Chain Rule). Thus $x \longmapsto F(g(x))$ is a primitive of $x \longmapsto f(g(x))g'(x)$ on any interval in the domain of $F \circ g$. So if we substitute $u = g(x)$, then we obtain a simpler integrand since

$$\int f(g(x))g'(x)\,dx = F(g(x)) = F(u) = \int f(u)\,du. \qquad (16)$$

This technique is worth trying if you can express the integrand in the form $f(g(x))g'(x)$, for some functions f and g, as illustrated in the next worked exercise.

Worked Exercise F31

Find a primitive of the function

$$x \longmapsto x^2(x^3+1)^8 \quad (x \in \mathbb{R}).$$

Solution

Differentiating x^3+1 gives $3x^2$, which suggests that we try putting $u = g(x) = x^3+1$.

Put $u = g(x) = x^3+1$. Then $\dfrac{du}{dx} = g'(x) = 3x^2$, so $du = 3x^2\,dx$.

Now substitute to express the integral in terms of u, adjusting by a multiplicative constant as necessary.

We have

$$\begin{aligned}\int x^2(x^3+1)^8\,dx &= \tfrac{1}{3}\int (x^3+1)^8\,3x^2\,dx\\ &= \tfrac{1}{3}\int u^8\,du.\end{aligned}$$

Next, evaluate the integral in terms of u.

$$\begin{aligned}&= \tfrac{1}{3}\times\tfrac{1}{9}u^9\\ &= \tfrac{1}{27}u^9\end{aligned}$$

Finally, substitute back for u to obtain the result in terms of x.

$$= \tfrac{1}{27}(x^3+1)^9.$$

Thus $\frac{1}{27}(x^3+1)^9$ is a primitive of $x^2(x^3+1)^8$.

You can always check your result by differentiating.

The following strategy summarises the approach taken in Worked Exercise F31.

Strategy F8

To find a primitive $\int f(g(x))g'(x)\,dx$ using integration by substitution, do the following.

1. Choose $u = g(x)$. Find $\dfrac{du}{dx} = g'(x)$ and hence express du in terms of x and dx.
2. Substitute $u = g(x)$ and replace $g'(x)\,dx$ by du (adjusting constants if necessary) to give $\int f(u)\,du$.
3. Find $\int f(u)\,du$.
4. Substitute $u = g(x)$ to give the required primitive in terms of x.

Remarks

1. If we are evaluating an integral, rather than finding a primitive, then there is no need to perform step 4 of Strategy F8. Instead, we can change the x-limits of integration into the corresponding u-limits:

$$\int_a^b f(g(x))g'(x)\,dx = \int_{g(a)}^{g(b)} f(u)\,du.$$

   For example, in Worked Exercise F31 we have

$$u = g(x) = x^3 + 1,$$

   so

$$\text{when } x = 0, \quad u = 1, \qquad \text{and when } x = 1, \quad u = 2.$$

   Hence

$$\int_0^1 x^2(x^3+1)^8\,dx = \tfrac{1}{3}\int_1^2 u^8\,du.$$

2. When applying equation (16), you may be able to spot a primitive F of f immediately, in which case you can write down the required primitive $F(g(x))$ directly without need for a substitution. For example, if you were able to spot the primitive in Worked Exercise F31, you could just write

$$\begin{aligned}\int x^2(x^3+1)^8\,dx &= \tfrac{1}{3}\int (x^3+1)^8\,3x^2\,dx\\ &= \tfrac{1}{3}\times\tfrac{1}{9}(x^3+1)^9 = \tfrac{1}{27}(x^3+1)^9.\end{aligned}$$

One particular situation in which this simplification can often be applied is with an integrand of the form $g'(x)/g(x)$ on an interval I, since we have

$$\int \frac{g'(x)}{g(x)}\,dx = \log(g(x)), \quad \text{if } g(x) > 0, \quad \text{for } x \in I, \tag{17}$$

because $\dfrac{d}{dx}\log(g(x)) = \dfrac{1}{g(x)} \times g'(x)$.

An example involving this useful formula appears in the next exercise.

Exercise F41

Find a primitive of each of the following functions.

(a) $f(x) = \sin(\sin 3x)\cos 3x \quad (x \in \mathbb{R})$

(b) $f(x) = x^2(2+3x^3)^7 \quad (x \in \mathbb{R})$

(c) $f(x) = x\sin(2x^2) \quad (x \in \mathbb{R})$

(d) $f(x) = x/(2+3x^2) \quad (x \in \mathbb{R})$

Exercise F42

Evaluate the integral

$$\int_0^1 \frac{e^x}{(1+e^x)^2}\,dx.$$

Our second substitution technique is a modification of the above method, which we call *backwards substitution*. It is based on the formula

$$\int f(x)\,dx = \int f(h(u))h'(u)\,du,$$

obtained from equation (16) by swapping the variables x and u. Here h is a function such that $x = h(u)$, often found by first writing $u = g(x)$ where g has an inverse function $g^{-1} = h$. Backwards substitution is worth trying if it makes part of the integrand significantly simpler. The next worked exercise gives an example of the use of this technique.

Worked Exercise F32

Find a primitive of the function

$$x \longmapsto \frac{e^{2x}}{(e^x - 1)^{1/2}} \quad (x \in (0, \infty)).$$

Solution

If $u = g(x) = (e^x - 1)^{1/2}$, then g has an inverse function $x = h(u) = \log(u^2 + 1)$.

Put $u = g(x) = (e^x - 1)^{1/2}$. Then the inverse function is $x = h(u) = \log(u^2 + 1)$, so $\dfrac{dx}{du} = h'(u) = \dfrac{2u}{u^2 + 1}$ and therefore $dx = \dfrac{2u}{u^2 + 1}\, du$.

Now substitute to express the integral in terms of u.

We have

$$\begin{aligned} \int \frac{e^{2x}}{(e^x - 1)^{1/2}}\, dx &= \int \frac{e^{\log(u^2+1)^2}}{u} \frac{2u}{u^2 + 1}\, du \\ &= \int \frac{(u^2 + 1)^2}{u} \frac{2u}{u^2 + 1}\, du \\ &= \int 2(u^2 + 1)\, du \end{aligned}$$

Next, evaluate the integral in terms of u.

$$= \tfrac{2}{3}u^3 + 2u$$

Finally, substitute back for u to obtain the result in terms of x.

$$= \tfrac{2}{3}(e^x - 1)^{3/2} + 2(e^x - 1)^{1/2}.$$

Thus $\frac{2}{3}(e^x - 1)^{3/2} + 2(e^x - 1)^{1/2}$ is a primitive of $\dfrac{e^{2x}}{(e^x - 1)^{1/2}}$.

The following strategy summarises the approach taken in Worked Exercise F32.

Strategy F9

To find a primitive $\displaystyle\int f(x)\,dx$ using integration by backwards substitution, do the following.

1. Choose $u = g(x)$, where g has an inverse function $x = h(u)$. Find $\dfrac{dx}{du} = h'(u)$ and hence express dx in terms of u and du.
2. Substitute $x = h(u)$ and replace dx by $h'(u)\,du$ to give a primitive in terms of u.
3. Find this primitive.
4. Substitute $u = g(x)$ to give the required primitive in terms of x.

As before, if we are evaluating an integral, then instead of step 4 in Strategy F9, we can change the x-limits of integration into the corresponding u-limits:

$$\int_a^b f(x)\,dx = \int_{g(a)}^{g(b)} f(g^{-1}(u))(g^{-1})'(u)\,du.$$

For example, in Worked Exercise F32 we have

$$u = g(x) = (e^x - 1)^{1/2},$$

so

$$\text{when } x = \log 2, \quad u = 1, \text{ and when } x = \log 3, \quad u = \sqrt{2}.$$

Hence

$$\int_{\log 2}^{\log 3} \frac{e^{2x}}{(e^x - 1)^{1/2}}\,dx = \int_1^{\sqrt{2}} 2(u^2 + 1)\,du.$$

Exercise F43

(a) Find a primitive of the function

$$f(x) = \frac{1}{3(x-1)^{3/2} + x(x-1)^{1/2}} \qquad (x \in (1, \infty)),$$

using the substitution $u = (x-1)^{1/2}$.

(b) Evaluate the integral

$$\int_0^{\log 3} e^x \sqrt{1 + e^x}\,dx.$$

Integration by parts

The technique of *integration by parts* is derived from the Product Rule for differentiation,

$$(fg)' = f'g + fg',$$

which implies that

$$\int fg' = fg - \int f'g, \quad \text{so} \quad \int_a^b fg' = [fg]_a^b - \int_a^b f'g.$$

This formula converts the problem of finding a primitive of fg' into the problem of finding a primitive of $f'g$. Integration by parts is worth trying if you can express the integrand as a product of two functions, $f(x)g'(x)$, where $f(x)$ becomes simpler on differentiation, and $g'(x)$ becomes not much more complicated on integration.

Here is a strategy for using integration by parts.

Strategy F10

To find a primitive $\int k(x)\,dx$ using integration by parts, do the following.

1. Write the original function k in the form fg', where f is a function that you can differentiate and g' is a function that you can integrate.
2. Use the formula $\int fg' = fg - \int f'g$.

The next two worked exercises give examples of using this technique.

Worked Exercise F33

Find

$$\int x \cos x \, dx.$$

Solution

We use integration by parts with $f(x) = x$ and $g'(x) = \cos x$.

Take $f(x) = x$ and $g'(x) = \cos x$, so that $f'(x) = 1$ and $g(x) = \sin x$. Then

$$\begin{aligned}\int x \cos x \, dx &= x \sin x - \int \sin x \, dx \\ &= x \sin x + \cos x.\end{aligned}$$

Sometimes we have to multiply the integrand by the factor 1 in order to be able to apply integration by parts, as in the following example.

Worked Exercise F34

Evaluate the integral

$$\int_0^1 \tan^{-1} x \, dx.$$

Solution

We use integration by parts, introducing the factor 1:

Here

$$f(x) = \tan^{-1} x, \quad g'(x) = 1,$$

so

$$f'(x) = \frac{1}{1+x^2}, \quad g(x) = x.$$

$$\begin{aligned}\int_0^1 \tan^{-1} x \, dx &= \int_0^1 1 \times \tan^{-1} x \, dx \\ &= \left[x \tan^{-1} x\right]_0^1 - \int_0^1 \frac{x}{1+x^2} \, dx \\ &= \tan^{-1} 1 - \left[\tfrac{1}{2}\log(1+x^2)\right]_0^1 \\ &= \frac{\pi}{4} - \left(\frac{1}{2}\log 2 - \frac{1}{2}\log 1\right) \\ &= \frac{\pi}{4} - \frac{1}{2}\log 2.\end{aligned}$$

Here we have used equation (17) with $g(x) = 1 + x^2$.

Exercise F44

(a) Find a primitive of the function

$$k(x) = x^{1/3} \log x \quad (x \in \mathbb{R}^+).$$

(b) Evaluate the integral

$$\int_0^{\pi/2} x^2 \cos x \, dx.$$

Hint: Use integration by parts twice.

Reduction formulas

Sometimes we need to evaluate an integral I_n that involves a non-negative integer n. A common approach to such integrals is to relate the value of I_n to the value of I_{n-1} or I_{n-2} by a **reduction formula** (sometimes called a *recurrence relation*) using integration by parts. Here is an example that will be important later in the unit.

Worked Exercise F35

Let

$$I_n = \int_0^{\pi/2} \sin^n x\, dx, \quad n = 0, 1, 2, \ldots .$$

(a) Evaluate I_0 and I_1.

(b) Prove that

$$I_n = \left(\frac{n-1}{n}\right) I_{n-2}, \quad \text{for } n \geq 2.$$

(c) Deduce the values of I_2, I_3, I_4 and I_5.

Solution

(a) We have $I_0 = \displaystyle\int_0^{\pi/2} 1\, dx = \pi/2$ and

$$I_1 = \int_0^{\pi/2} \sin x\, dx = \big[-\cos x\big]_0^{\pi/2} = 1.$$

(b) We write

$$I_n = \int_0^{\pi/2} \sin x\, \sin^{n-1} x\, dx.$$

We integrate $\sin x$ and differentiate $\sin^{n-1} x$.

Using integration by parts we find that, for $n \geq 2$,

$$\begin{aligned} I_n &= \big[(-\cos x)\sin^{n-1} x\big]_0^{\pi/2} - \\ &\qquad \int_0^{\pi/2} (-\cos x)(n-1)\sin^{n-2} x \cos x\, dx \\ &= 0 + (n-1)\int_0^{\pi/2} \cos^2 x \sin^{n-2} x\, dx \\ &= (n-1)\int_0^{\pi/2} (1-\sin^2 x)\sin^{n-2} x\, dx \\ &= (n-1)\left(\int_0^{\pi/2} \sin^{n-2} x\, dx - \int_0^{\pi/2} \sin^n x\, dx\right) \\ &= (n-1)(I_{n-2} - I_n). \end{aligned}$$

We can rearrange this equation to give

$$nI_n = (n-1)I_{n-2}, \quad \text{so} \quad I_n = \left(\frac{n-1}{n}\right) I_{n-2}, \quad \text{for } n \geq 2.$$

(c) Using the result of part (b) with $n = 2, 3, 4, 5$ in turn, we obtain

$$\begin{aligned}
I_2 &= \frac{1}{2} I_0 = \frac{1}{2} \cdot \frac{\pi}{2} = \frac{\pi}{4}, \\
I_3 &= \frac{2}{3} I_1 = \frac{2}{3}, \\
I_4 &= \frac{3}{4} I_2 = \frac{3}{4} \cdot \frac{\pi}{4} = \frac{3\pi}{16}, \\
I_5 &= \frac{4}{5} I_3 = \frac{4}{5} \cdot \frac{2}{3} = \frac{8}{15}.
\end{aligned}$$

By repeatedly applying the reduction formula in Worked Exercise F35, we obtain the general formulas

$$I_{2n} = \frac{1}{2} \cdot \frac{3}{4} \cdot \frac{5}{6} \cdot \cdots \cdot \frac{2n-1}{2n} \cdot \frac{\pi}{2}$$

and

$$I_{2n+1} = \frac{2}{3} \cdot \frac{4}{5} \cdot \frac{6}{7} \cdot \cdots \cdot \frac{2n}{2n+1}.$$

We will use these formulas in Subsection 3.2.

Exercise F45

Let

$$I_n = \int_0^1 e^x x^n \, dx, \quad n = 0, 1, 2, \ldots.$$

(a) Evaluate I_0.

(b) Prove that

$$I_n = e - nI_{n-1}, \quad \text{for } n = 1, 2, \ldots.$$

(c) Deduce the values of I_1, I_2, I_3 and I_4.

3 Inequalities, sequences and series

Often it is not possible to evaluate an integral explicitly, and a numerical estimate for its value is sufficient. This situation can arise both in applications of mathematics and in proofs that involve integration. In this section we study some inequalities satisfied by integrals, and apply these to find two remarkable formulas for π, and to decide whether certain series are convergent or divergent.

3.1 Inequalities for integrals

The basic inequality rules for integrals are as follows.

Theorem F57 Inequality Rules

Let f and g be integrable on $[a, b]$.

(a) If $f(x) \leq g(x)$, for $x \in [a, b]$, then

$$\int_a^b f \leq \int_a^b g.$$

(b) If $m \leq f(x) \leq M$, for $x \in [a, b]$, then

$$m(b-a) \leq \int_a^b f \leq M(b-a).$$

Proof The proofs of parts (a) and (b) of Theorem F57 are illustrated in Figure 15.

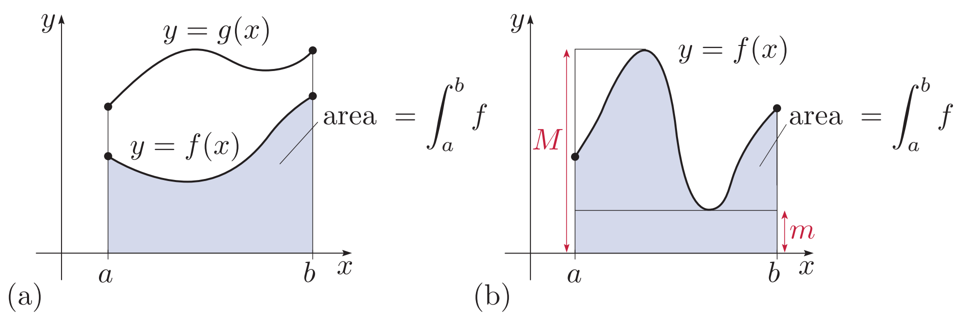


Figure 15 The proofs of the Inequality Rules parts (a) and (b)

(a) Let P be any partition of $[a, b]$. Since

$$f(x) \leq g(x), \quad \text{for } x \in [a, b],$$

the infimum of f on each subinterval of P is less than or equal to the infimum of g on that subinterval, and therefore $L(f, P) \leq L(g, P)$. Thus

$$\int_a^b f = \sup_P L(f, P) \leq \sup_P L(g, P) = \int_a^b g,$$

since f and g are both integrable on $[a, b]$.

(b) Since $f(x) \leq M$ for $x \in [a, b]$, it follows from part (a), with $g(x) = M$, that

$$\int_a^b f \leq \int_a^b M\,dx = M(b-a).$$

The proof of the left-hand inequality is similar. ■

The Inequality Rules allow us to estimate a complicated integral by evaluating a simpler one, as in the next worked exercise.

Worked Exercise F36

Prove the following inequalities.

(a) $\displaystyle\int_0^1 \frac{x^3}{2-\sin^4 x}\,dx \leq \tfrac{1}{4}\log 2$ (b) $\displaystyle\frac{3}{\sqrt{34}} \leq \int_{-1}^2 \frac{dx}{\sqrt{2+x^5}} \leq 3$

Solution

(a) Since

$$|\sin x| \leq |x|, \quad \text{for } x \in \mathbb{R},$$

This is Corollary D46 from Subsection 2.3 of Unit D4.

it follows that

$$\sin^4 x \leq x^4, \quad \text{for } x \in \mathbb{R}.$$

Thus $2 - \sin^4 x \geq 2 - x^4 > 0$ for $x \in [0,1]$, so

$$\frac{x^3}{2-\sin^4 x} \leq \frac{x^3}{2-x^4}, \quad \text{for } x \in [0,1].$$

Hence, by Inequality Rule (a), we have

$$\begin{aligned}\int_0^1 \frac{x^3}{2-\sin^4 x}\,dx &\leq \int_0^1 \frac{x^3}{2-x^4}\,dx \\ &= \left[-\tfrac{1}{4}\log(2-x^4)\right]_0^1 \\ &= -\tfrac{1}{4}(\log 1 - \log 2) \\ &= \tfrac{1}{4}\log 2.\end{aligned}$$

Here we have used equation (17) with $g(x) = 2 - x^4$.

(b) Since the function $x \longmapsto \sqrt{2+x^5}$ is increasing on $[-1,2]$, we have

$$1 \leq \sqrt{2+x^5} \leq \sqrt{34}, \quad \text{for } x \in [-1,2],$$

so

$$\frac{1}{\sqrt{34}} \leq \frac{1}{\sqrt{2+x^5}} \leq 1, \quad \text{for } x \in [-1,2].$$

Since the length of the interval $[-1,2]$ is 3, it follows from Inequality Rule (b) that

$$\frac{3}{\sqrt{34}} \leq \int_{-1}^2 \frac{dx}{\sqrt{2+x^5}} \leq 3.$$

Exercise F46

Use the Inequality Rules to prove the following inequalities.

(a) $\displaystyle\int_1^3 x\sin\big(1/x^{10}\big)\,dx \leq 4$ (b) $\frac{1}{2} \leq \displaystyle\int_0^{1/2} e^{x^2}\,dx \leq \frac{1}{2}e^{1/4}$

We saw in Subsection 1.3 that if the function f is integrable on $[a,b]$, then so is $|f|$. We now use the Inequality Rules to obtain an inequality involving the integrals of f and $|f|$, known as the Triangle Inequality for integrals. The name arises because of the similarity between this inequality and the Triangle Inequality for numbers:

$$\left|\sum_{i=1}^{n} a_i\right| \leq \sum_{i=1}^{n} |a_i|.$$

Theorem F58 Triangle Inequality for integrals

Let f be integrable on $[a,b]$. Then

$$\left|\int_a^b f\right| \leq \int_a^b |f|.$$

Furthermore, if $|f(x)| \leq M$ for $x \in [a,b]$, then

$$\left|\int_a^b f\right| \leq M(b-a).$$

Proof We know that, for all $x \in [a,b]$,

$$-|f(x)| \leq f(x) \leq |f(x)|.$$

Since $|f|$ is integrable on $[a,b]$, it follows from Inequality Rule (a) that

$$-\int_a^b |f| \leq \int_a^b f \leq \int_a^b |f|,$$

which is equivalent to

$$\left|\int_a^b f\right| \leq \int_a^b |f|.$$

Finally, if $|f(x)| \leq M$ for $x \in [a,b]$, then, by the above inequality and Inequality Rule (b),

$$\left|\int_a^b f\right| \leq \int_a^b |f| \leq M(b-a).$$ ■

Sometimes we use the Inequality Rules and the Triangle Inequality in combination, as in the next worked exercise.

Worked Exercise F37

Prove that

$$\left|\int_0^{\pi/2} \frac{x-\pi/2}{2+\cos x}\,dx\right| \leq \frac{\pi^2}{16}.$$

Solution

By the Triangle Inequality for integrals,

$$\begin{aligned}\left|\int_0^{\pi/2} \frac{x-\pi/2}{2+\cos x}\,dx\right| &\leq \int_0^{\pi/2} \left|\frac{x-\pi/2}{2+\cos x}\right| dx \\ &= \int_0^{\pi/2} \frac{\pi/2-x}{2+\cos x}\,dx. \qquad (*)\end{aligned}$$

For $0 \leq x \leq \pi/2$, we have $\pi/2 - x \geq 0$ and $\cos x \geq 0$.

Next, since

$$2+\cos x \geq 2, \quad \text{for } x \in [0, \pi/2],$$

we have

$$\frac{1}{2+\cos x} \leq \frac{1}{2}, \quad \text{for } x \in [0, \pi/2].$$

Thus, by Inequality Rule (a) and statement $(*)$,

$$\begin{aligned}\left|\int_0^{\pi/2} \frac{x-\pi/2}{2+\cos x}\,dx\right| &\leq \int_0^{\pi/2} \frac{1}{2}\left(\frac{\pi}{2}-x\right) dx \\ &= \frac{1}{2}\left[\frac{\pi}{2}x - \frac{1}{2}x^2\right]_0^{\pi/2} \\ &= \frac{1}{2}\left(\frac{\pi^2}{4} - \frac{\pi^2}{8}\right) \\ &= \frac{\pi^2}{16}.\end{aligned}$$

Exercise F47

Prove the following inequalities.

(a) $\left|\int_1^4 \frac{\sin(1/x)}{2+\cos(1/x)}\,dx\right| \leq 3$ (b) $\left|\int_0^{\pi/4} \frac{\tan x}{3-\sin(x^2)}\,dx\right| \leq \frac{1}{4}\log 2$

3.2 Wallis' Formula

In Worked Exercise F35 we used a reduction formula to show that if

$$I_n = \int_0^{\pi/2} \sin^n x \, dx, \quad n = 0, 1, 2, \dots,$$

then

$$I_0 = \frac{\pi}{2}, \quad I_1 = 1 \quad \text{and} \quad I_n = \left(\frac{n-1}{n}\right) I_{n-2}, \quad \text{for } n \geq 2. \tag{18}$$

We also remarked that by repeatedly applying equations (18) we obtain

$$I_{2n} = \frac{1}{2} \cdot \frac{3}{4} \cdot \frac{5}{6} \cdot \dots \cdot \frac{2n-1}{2n} \cdot \frac{\pi}{2}, \quad \text{for } n \geq 1, \tag{19}$$

and

$$I_{2n+1} = \frac{2}{3} \cdot \frac{4}{5} \cdot \frac{6}{7} \cdot \dots \cdot \frac{2n}{2n+1}, \quad \text{for } n \geq 1. \tag{20}$$

Note that the formula for I_{2n} involves π, but the formula for I_{2n+1} does not.

We now use these results, together with inequalities between various integrals of the form I_n, to establish two remarkable formulas for π, the first of which is Wallis' Formula.

Theorem F59

(a) $$\lim_{n\to\infty} \left(\frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5} \cdot \frac{6}{5} \cdot \frac{6}{7} \cdot \dots \cdot \frac{2n}{2n-1} \cdot \frac{2n}{2n+1} \right) = \frac{\pi}{2}$$

(b) $$\lim_{n\to\infty} \frac{(n!)^2\, 2^{2n}}{(2n)!\sqrt{n}} = \sqrt{\pi}$$

John Wallis

John Wallis (1616–1703) was the most influential English mathematician before the rise of Isaac Newton. He was a cryptographer to the Parliamentarians during the Civil War and in 1649 was appointed Savilian Professor of Geometry at Oxford.

His most important work is his *Arithmetica Infinitorum* (Arithmetic of Infinitesimals), published in 1656. It was in this work that Wallis derived the formula which bears his name, and it was through studying this work that Newton came to discover his version of the binomial theorem.

Wallis' treatise on conic sections, published in 1655, contains the first publication of the symbol for infinity, ∞.

Part (c) of the next exercise gives an identity needed in the proof of Theorem F59. If you are short of time, then you may wish to skip this part of the exercise, which is quite challenging, and skim read the solution and also the proof of the theorem.

Exercise F48

For $n = 1, 2, \dots,$ let

$$a_n = \frac{2}{1}\cdot\frac{2}{3}\cdot\frac{4}{3}\cdot\frac{4}{5}\cdot\frac{6}{5}\cdot\frac{6}{7}\cdot\cdots\cdot\frac{2n}{2n-1}\cdot\frac{2n}{2n+1} \quad \text{and} \quad b_n = \frac{(n!)^2\,2^{2n}}{(2n)!\sqrt{n}}.$$

(a) Evaluate a_n and b_n, for $n = 1, 2, 3$.

(b) Verify that

$$b_n^2 = \left(\frac{2n+1}{n}\right)a_n, \quad \text{for } n = 1, 2, 3.$$

(c) Prove that

$$b_n^2 = \left(\frac{2n+1}{n}\right)a_n, \quad \text{for } n = 1, 2, \dots.$$

Proof of Theorem F59

Let

$$I_n = \int_0^{\pi/2} \sin^n x\,dx, \quad n = 0, 1, 2, \dots,$$

and let the sequences (a_n) and (b_n) be as given in Exercise F48.

(a) Using equations (19) and (20), we obtain, for $n \geq 1$,

$$\frac{I_{2n}}{I_{2n+1}} = \frac{1\cdot 3\cdot 3\cdot 5\cdot\ \cdots\ \cdot(2n-1)(2n+1)}{2\cdot 2\cdot 4\cdot 4\cdot\ \cdots\ \cdot(2n)(2n)}\,\frac{\pi}{2} = \frac{1}{a_n}\cdot\frac{\pi}{2},$$

so

$$a_n = \left(\frac{I_{2n+1}}{I_{2n}}\right)\frac{\pi}{2}.$$

Thus to prove part (a), it is sufficient to show that

$$\frac{I_{2n+1}}{I_{2n}} \to 1 \ \text{ as } n \to \infty. \tag{21}$$

We do this as follows. Since $0 \leq \sin x \leq 1$, for $x \in [0, \pi/2]$, we have

$$\sin^{2n} x \geq \sin^{2n+1} x \geq \sin^{2n+2} x, \quad \text{for } x \in [0, \pi/2].$$

It follows by Inequality Rule (a) that

$$I_{2n} \geq I_{2n+1} \geq I_{2n+2}.$$

Thus, on dividing by I_{2n}, we obtain

$$1 \geq \frac{I_{2n+1}}{I_{2n}} \geq \frac{I_{2n+2}}{I_{2n}} = \frac{2n+1}{2n+2} = \frac{2+1/n}{2+2/n},$$

by equation (18). On taking the limit as $n \to \infty$, we deduce that statement (21) holds, by the Squeeze Rule for sequences from Unit D2.

(b) We know, from Exercise F48(c), that

$$b_n^2 = \left(\frac{2n+1}{n}\right) a_n = \left(2 + \frac{1}{n}\right) a_n.$$

By part (a),

$$a_n \to \frac{\pi}{2} \quad \text{as } n \to \infty,$$

so, by the Product Rule for sequences,

$$b_n^2 \to 2 \times \frac{\pi}{2} = \pi \quad \text{as } n \to \infty.$$

Hence, by the continuity of the square root function,

$$b_n \to \sqrt{\pi} \quad \text{as } n \to \infty.$$

■

3.3 The Integral Test

In this subsection we introduce a method based on integration for determining the convergence or divergence of certain series of the form $\sum_{n=1}^{\infty} f(n)$, where the function f is positive and decreasing and tends to 0. The method is based on the fact that it is often easier to evaluate an integral than a sum which has a similar behaviour.

Theorem F60 Integral Test

Let the function f be positive and decreasing on $[1, \infty)$, and suppose that $f(x) \to 0$ as $x \to \infty$. Then

(a) $\displaystyle\sum_{n=1}^{\infty} f(n)$ converges if the sequence $\displaystyle\left(\int_1^n f\right)$ is bounded above

(b) $\displaystyle\sum_{n=1}^{\infty} f(n)$ diverges if $\displaystyle\int_1^n f \to \infty$ as $n \to \infty$.

Remarks

1. The Integral Test is also called the *Maclaurin Integral Test*.
2. In both parts (a) and (b), the number 1 can be replaced by any positive integer.

Colin Maclaurin (1698–1746) was a Scottish mathematician who spent most of his career as professor of mathematics at Edinburgh University, having been appointed on the recommendation of Isaac Newton. A popular teacher, he was described as a 'favourite professor' and the 'life and soul' of the university. He is notable for having extended Newton's work on the calculus and geometry. His two-volume *Treatise of Fluxions* (1742), which was the first systematic treatment of Newton's methods, contains a detailed discussion of infinite series and includes the Integral Test in verbal form.

The Integral Test was rediscovered by Cauchy – he published it in 1827 – and consequently it is also known as the Maclaurin–Cauchy Integral Test.

Colin Maclaurin

Proof of Theorem F60 For $n = 2, 3, \dots,$ let $s_n = f(1) + f(2) + \cdots + f(n)$ be the nth partial sum of the series $\sum_{n=1}^{\infty} f(n)$, and let P_{n-1} be the standard partition of $[1, n]$ with $n - 1$ subintervals of length 1, that is,

$$\{[1,2], \dots, [i, i+1], \dots, [n-1, n]\}.$$

Since f is decreasing on $[1, \infty)$, we have, for $i = 1, 2, \dots, n-1$,

$$m_i = f(i+1) \quad \text{and} \quad M_i = f(i).$$

This is illustrated in Figure 16.

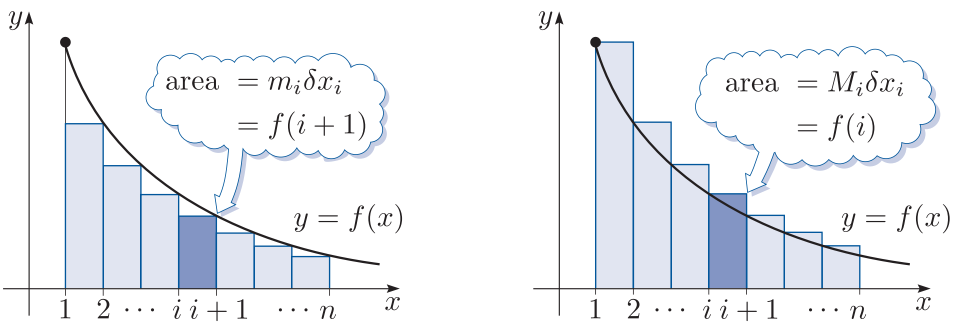


Figure 16 The lower and upper Riemann sums of f

Also, each subinterval in the partition has length 1. Hence the lower and upper Riemann sums for f on $[1, n]$ are

$$L(f, P_{n-1}) = \sum_{i=1}^{n-1} m_i \times 1 = f(2) + \cdots + f(n) = s_n - f(1)$$

and

$$U(f, P_{n-1}) = \sum_{i=1}^{n-1} M_i \times 1 = f(1) + \cdots + f(n-1) = s_n - f(n).$$

Since f is bounded and monotonic on $[1,n]$, it follows from Theorem F48 in Subsection 1.2 that the integral $I_n = \displaystyle\int_1^n f$ exists and satisfies

$$L(f, P_{n-1}) \leq I_n \leq U(f, P_{n-1}),$$

so

$$s_n - f(1) \leq I_n \leq s_n - f(n). \tag{22}$$

We now consider two cases separately.

Case (a): The sequence of integrals (I_n) is bounded above.

In this case, there exists $M \in \mathbb{R}$ such that

$$I_n \leq M, \quad \text{for } n = 2, 3, \ldots.$$

It then follows from the left-hand inequality in statement (22) that

$$s_n \leq f(1) + M, \quad \text{for } n = 2, 3, \ldots.$$

Thus the increasing sequence (s_n) is bounded above, so it is convergent, by the Monotone Convergence Theorem (Theorem D22 in Subsection 5.1 of Unit D2).

Hence the series $\sum_{n=1}^{\infty} f(n)$ is convergent.

Case (b): The sequence (I_n) is not bounded above.

The sequence (I_n) is increasing, since

$$I_{n+1} - I_n = \int_n^{n+1} f \geq 0,$$

so in this case

$$I_n \to \infty \text{ as } n \to \infty.$$

It follows from the right-hand inequality in statement (22) that

$$s_n \geq I_n, \quad \text{for } n = 2, 3, \ldots.$$

Thus, by the Squeeze Rule for sequences which tend to infinity (Theorem D18 in Subsection 4.3 of Unit D2),

$$s_n \to \infty \text{ as } n \to \infty.$$

Hence the series $\sum_{n=1}^{\infty} f(n)$ is divergent. ■

In Theorem D33 of Unit D3 *Series* you saw that the basic series $\sum_{n=1}^{\infty} 1/n^p$ converges for $p \geq 2$ and diverges for $0 < p \leq 1$. We can now use the Integral Test to deduce the behaviour of this series for all $p > 0$. We first consider the case when $p = 1$.

Worked Exercise F38

Use the fact that $\displaystyle\int \frac{dx}{x} = \log x$ to prove that $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n}$ is divergent.

Solution

Let

$$f(x) = \frac{1}{x} \quad (x \in [1, \infty)).$$

Then f is positive and decreasing on $[1, \infty)$, and

$$f(x) \to 0 \ \text{ as } x \to \infty.$$

Also, for $n \geq 1$,

$$\begin{aligned} \int_1^n f &= \int_1^n \frac{dx}{x} \\ &= \big[\log x\big]_1^n \\ &= \log n \to \infty \ \text{ as } n \to \infty. \end{aligned}$$

Hence, by part (b) of the Integral Test, the series diverges.

We now consider the series $\sum_{n=1}^{\infty} 1/n^p$ when $p > 0$ but $p \neq 1$.

Worked Exercise F39

Use the Integral Test to determine the behaviour of the series

$$\sum_{n=1}^{\infty} \frac{1}{n^p}, \quad \text{for } p > 0,\ p \neq 1.$$

Solution

Let $p > 0$ and $p \neq 1$, and let

$$f(x) = 1/x^p \quad (x \in [1, \infty)).$$

Then f is positive and decreasing on $[1, \infty)$, and

$$f(x) \to 0 \ \text{ as } x \to \infty.$$

Also, for $n \in \mathbb{N}$,

$$\int_1^n f = \int_1^n \frac{dx}{x^p} = \left[\frac{x^{1-p}}{1-p}\right]_1^n = \frac{n^{1-p} - 1}{1-p}. \qquad (*)$$

First suppose that $p > 1$. Then $p - 1 > 0$, so equation $(*)$ gives

$$\int_1^n f = \frac{1}{p-1}\left(1 - \frac{1}{n^{p-1}}\right) < \frac{1}{p-1}.$$

Hence the sequence of integrals $\left(\int_1^n f\right)$ is bounded above, so it follows from part (a) of the Integral Test that the series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges.

Now suppose that $0 < p < 1$. Then $1 - p > 0$, so $1/n^{1-p} \to 0$ as $n \to \infty$, and hence

$$n^{1-p} \to \infty \quad \text{as } n \to \infty,$$

by the Reciprocal Rule for sequences (Theorem D16 from Subsection 4.1 of Unit D2). We deduce from equation $(*)$ that

$$\int_1^n f \to \infty \quad \text{as } n \to \infty.$$

Hence, by part (b) of the Integral Test, the series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ diverges.

So, the series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges for $p > 1$, and diverges for $0 < p < 1$.

Exercise F49

Show that

$$\int \frac{dx}{x(\log x)^2} = -\frac{1}{\log x},$$

and hence prove that

$$\sum_{n=2}^{\infty} \frac{1}{n(\log n)^2} \text{ is convergent.}$$

4 Stirling's Formula

This section concerns the value of the quantity $n!$ which arises in many problems in probability. You will see that integration techniques give an excellent estimate for $n!$, called Stirling's Formula, which can be expressed as

$$\lim_{n\to\infty} \frac{n!}{\sqrt{2\pi n}\,(n/e)^n} = 1.$$

4.1 Comparing functions of n

In this subsection we define a relation $\sim$ on the set of all positive functions with domain $\mathbb{N}$ to provide a way of comparing the behaviour of such functions for large values of n. This relation is, in fact, an equivalence relation, although we do not prove this here or make explicit use of these properties. (You met equivalence relations in Unit A3 *Mathematical language and proof* and you may like to try to prove that $\sim$ is an equivalence relation yourself; it is included as an exercise in the additional exercise booklet for this unit.)

Definition

For positive functions f and g with domain $\mathbb{N}$, we write

$$f(n) \sim g(n) \quad \text{as } n \to \infty$$

to mean

$$\frac{f(n)}{g(n)} \to 1 \quad \text{as } n \to \infty.$$

For example,

$$n^2 + n \sim n^2 \quad \text{as } n \to \infty,$$

since $n^2 + n > 0$ and $n^2 > 0$, for $n = 1, 2, \ldots$, and

$$\frac{n^2 + n}{n^2} \to 1 \quad \text{as } n \to \infty.$$

Note that the statement

$$f(n) \sim g(n) \quad \text{as } n \to \infty$$

does *not* imply that $f(n) - g(n)$ tends to zero or is even bounded. For instance, in the above example, $(n^2 + n) - n^2 = n$ tends to infinity. Note that often we omit 'as $n \to \infty$' when writing expressions of this type.

We have the following Combination Rules for $\sim$. These rules follow from the Combination Rules for sequences; we omit the proofs.

Theorem F61 Combination Rules for $\sim$

If $f_1(n) \sim g_1(n)$ and $f_2(n) \sim g_2(n)$, then:

Sum Rule $f_1(n) + f_2(n) \sim g_1(n) + g_2(n)$

Multiple Rule $\lambda f_1(n) \sim \lambda g_1(n)$, for $\lambda \in \mathbb{R}^+$

Product Rule $f_1(n) f_2(n) \sim g_1(n) g_2(n)$

Quotient Rule $\dfrac{f_1(n)}{f_2(n)} \sim \dfrac{g_1(n)}{g_2(n)}$.

Note that if $l > 0$, then the statements

$$f(n) \to l \quad \text{as } n \to \infty$$

and

$$f(n) \sim l \quad \text{as } n \to \infty$$

are equivalent, since each is equivalent to the statement

$$\frac{f(n)}{l} \to 1 \quad \text{as } n \to \infty.$$

4.2 Calculating factorials

For small values of n we can evaluate $n!$ directly by multiplication or by using a scientific calculator.

Exercise F50

Complete the following table of values of $n!$.

n	$n!$	n	$n!$	n	$n!$
1	1	6	720	20	$2.432\ldots \times 10^{18}$
2	2	7	5 040	30	
3	6	8	40 320	40	
4	24	9	362 880	50	
5	120	10	3 628 800	60	

As n increases, $n!$ grows very quickly, and is soon beyond the range of a calculator. Many calculations in probability theory involve $n!$ for large values of n, so it is important to be able to estimate this quantity as accurately as possible. The following result gives us a way of doing this.

Theorem F62 Stirling's Formula

$$n! \sim \sqrt{2\pi n}\,(n/e)^n \quad \text{as } n \to \infty.$$

We give the proof of this result in Subsection 4.3.

James Stirling (1692–1770) was a Scottish mathematician who made many contributions to analysis. He was friends with Isaac Newton and Colin Maclaurin, and at Maclaurin's request he provided help and criticisms when Maclaurin's *Treatise of Fluxions* was in proof.

Stirling's most important work was his *Methodus Differentialis*, published in 1730. It focuses largely on infinite series and was directly stimulated by earlier work of Wallis and Newton. It includes a series for $\log n!$, now usually written in the form known as Stirling's Formula, given above. Earlier in 1730, Abraham de Moivre (1667–1754), famous for his work on the laws of chance, had published a similar formula for $\log n!$, but he had not been able to find a precise value for the constant which Stirling showed to be $\sqrt{2\pi}$.

Exercise F51

Use a calculator to evaluate $\sqrt{2\pi n}\,(n/e)^n$ for the following values of n.

(a) $n = 5$ (b) $n = 10$ (c) $n = 50$

As you saw in Exercise F51, even for small values of n, Stirling's Formula gives reasonable approximations to $n!$, and the relative error (that is, the error expressed as a percentage of the true value) decreases as n increases.

n	$n!$	Stirling's approximation	Relative error
10	3 628 800	3 598 696	0.83%
20	2.433×10^{18}	2.423×10^{18}	0.42%
50	3.041×10^{64}	3.036×10^{64}	0.16%
100	9.333×10^{157}	9.325×10^{157}	0.08%

In fact it can be shown by a more careful argument that

$$e^{1/(12n+1)} \leq \frac{n!}{\sqrt{2\pi n}\,(n/e)^n} \leq e^{1/(12n)}, \quad \text{for } n \geq 1.$$

For example, if $n = 10$, then

$$e^{1/(12n+1)} = e^{1/121} = 1.008\,29\ldots \quad \text{and} \quad e^{1/(12n)} = e^{1/120} = 1.008\,36\ldots,$$

which indicates a relative error of about 0.8%, as shown in the above table.

Stirling's Formula can be used to give estimates of probabilities. For example, if we toss n fair coins, then it can be shown that the probability of obtaining exactly r heads is $\binom{n}{r}\frac{1}{2^n}$, where $\binom{n}{r} = \frac{n!}{r!(n-r)!}$ is the *binomial coefficient* which you met in Subsection 3.4 of Unit D1 (we do not prove this here). Therefore, if we toss 200 coins, then the probability of obtaining exactly 100 heads and 100 tails is

$$\begin{aligned}
\binom{200}{100}\frac{1}{2^{200}} &= \frac{200!}{(100!)^2\,2^{200}} \\
&\approx \frac{\sqrt{400\pi}\,(200/e)^{200}}{\left(\sqrt{200\pi}\,(100/e)^{100}\right)^2 2^{200}} \\
&= \frac{\sqrt{400\pi}\,\left(200^{200}/e^{200}\right)}{200\pi\,(100^{200}/e^{200})\,2^{200}} \\
&= \frac{\sqrt{400\pi}}{200\pi} \\
&= \frac{1}{10\sqrt{\pi}} \\
&= \frac{1}{17.724\ldots},
\end{aligned}$$

perhaps rather higher than you might expect.

The following exercises give you a chance to practise using Stirling's Formula. In some of these exercises you will need to use the Combination Rules for $\sim$ from Subsection 4.1.

Exercise F52

Use Stirling's Formula to estimate each of the following numbers (giving your answers to two significant figures).

(a) $\binom{300}{150}\frac{1}{2^{300}}$ (b) $\frac{300!}{(100!)^3}\frac{1}{3^{300}}$

Exercise F53

Use Stirling's Formula to determine a number λ such that

$$\binom{4n}{2n}\Big/\binom{2n}{n} \sim \lambda 2^{2n} \quad \text{as } n \to \infty.$$

Exercise F54

Use Stirling's Formula to prove that

$$\lim_{n\to\infty}\left(\frac{n^n}{n!}\right)^{1/n} = e.$$

Hint: You can assume that if $f(n) \sim g(n)$, then $(f(n))^{1/n} \sim (g(n))^{1/n}$. (This result about $\sim$ holds because if $n \in \mathbb{N}$, then

$$a \leq a^{1/n} < 1, \quad \text{for } 0 < a < 1,$$

and

$$1 < a^{1/n} \leq a, \quad \text{for } a > 1,$$

by the rules for inequalities.)

4.3 Proof of Stirling's Formula (optional)

The proof of Stirling's Formula is quite long. If you don't have time to study it properly, then you may find it interesting to skim through this subsection and see how integration is used to obtain an estimate of this type.

The idea of the proof is to consider the graph of $y = \log x$ and a sequence of small sets which lie below $y = \log x$ and above the graph of a polygonal approximation to $y = \log x$. We show that the areas of these small sets form a convergent series and deduce from this that the sequence

$$a_n = \frac{n^{n+(1/2)}}{e^{n-1}n!}, \quad n = 2, 3, \ldots,$$

is convergent. Finally we use Theorem F59(b) to find the limit of the sequence (a_n), and this gives Stirling's Formula.

Proof of Theorem F62 We need to show that

$$n! \sim \sqrt{2\pi n}\,(n/e)^n \text{ as } n \to \infty.$$

We begin the proof by considering the function

$$f(x) = \log x.$$

For $n = 2, 3, \ldots$, let P_{n-1} be the standard partition of the interval $[1, n]$ with $n - 1$ subintervals:

$$\{[1, 2], \ldots, [i, i + 1], \ldots, [n - 1, n]\}.$$

Now consider the sequence $(c_n)_2^\infty$, where c_n is the total area of the segments which lie between the graph

$$y = \log x, \quad x \in [1, n],$$

and the polygonal graph with vertices

$$(1, 0), \quad (2, \log 2), \quad (3, \log 3), \quad \ldots, \quad (n, \log n),$$

as illustrated in Figure 17. This set consists of $n-1$ thin segments.

The function f is concave (that is, its derivative is decreasing), so the line segment joining

$$(i, \log i) \text{ to } (i+1, \log(i+1))$$

lies below the graph $y = \log x$.

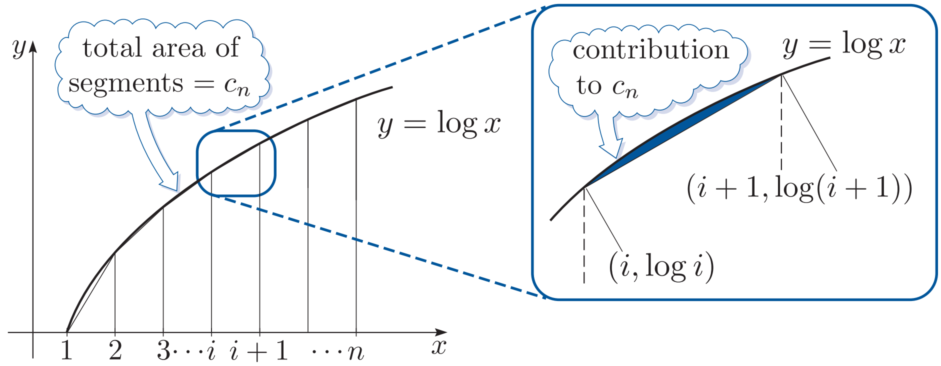


Figure 17 An approximation for the area under the graph of $\log x$

The area of the set between $y = \log x$, $x \in [1, n]$, and the x-axis is

$$\begin{aligned} \int_1^n \log x \, dx &= \left[x \log x - x\right]_1^n \\ &= n \log n - (n-1). \end{aligned} \tag{23}$$

The area between the polygonal graph and the x-axis is

$$\tfrac{1}{2}(L(f, P_{n-1}) + U(f, P_{n-1})). \tag{24}$$

This is because, on each subinterval $[i, i+1]$, the function f attains its minimum when $x = i$ and its maximum when $x = i+1$, so the area of the trapezium whose base is the subinterval is equal to the average of these two numbers.

Since f is increasing, we have

$$\begin{aligned} L(f, P_{n-1}) &= \log 1 + \log 2 + \cdots + \log(n-1) \\ &= \log(n-1)! \\ &= \log(n!/n) \\ &= \log n! - \log n \end{aligned} \tag{25}$$

and

$$\begin{aligned} U(f, P_{n-1}) &= \log 2 + \log 3 + \cdots + \log n \\ &= \log n!. \end{aligned} \tag{26}$$

Substituting from equations (25) and (26) into equation (24), we find that the area between the polygonal graph and the x-axis is

$$\tfrac{1}{2}(\log n! - \log n + \log n!) = \log n! - \tfrac{1}{2}\log n. \tag{27}$$

It follows from equations (23) and (27) that

$$\begin{aligned} c_n &= n\log n - (n-1) - \log n! + \tfrac{1}{2}\log n \\ &= \log\left(\frac{n^{n+(1/2)}}{e^{n-1}n!}\right). \end{aligned}$$

The sequence (c_n) is positive and increasing. Also,

$$c_n \leq \log 2, \quad \text{for } n \in \mathbb{N},$$

since the $n-1$ segments which contribute to the area c_n can be translated so that they all lie (without overlapping) in the triangle with vertices $(1,0)$, $(1,\log 2)$ and $(2,\log 2)$, as illustrated in Figure 18. This geometric property holds because the function log is concave. It follows from the Monotone Convergence Theorem (Theorem D22 in Subsection 5.1 of Unit D2) that the sequence (c_n) is convergent.

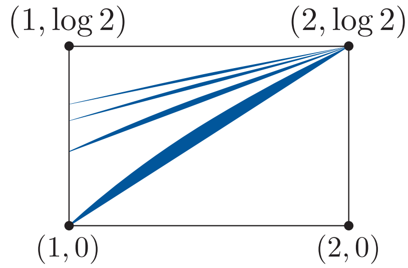


Figure 18 The segments of area c_n

Next, since the exponential function is continuous, the sequence

$$a_n = e^{c_n}, \quad n = 2, 3, \ldots,$$

is convergent also. Thus

$$a_n = \frac{n^{n+(1/2)}}{e^{n-1}n!} \to L \text{ as } n \to \infty, \tag{28}$$

for some non-zero number L.

To find L, we consider the quotient

$$\frac{a_n^2}{a_{2n}} = \frac{n^{2n+1}}{e^{2n-2}(n!)^2} \bigg/ \frac{(2n)^{2n+(1/2)}}{e^{2n-1}(2n)!} = \frac{(2n)!\, n^{1/2}}{(n!)^2\, 2^{2n}} \times \frac{e}{\sqrt{2}}.$$

We now let $n \to \infty$ in this equation. We have

$$a_n \to L, \quad a_{2n} \to L \quad \text{and} \quad \frac{(2n)!\, n^{1/2}}{(n!)^2\, 2^{2n}} \to \frac{1}{\sqrt{\pi}},$$

by Theorem F59(b).

Recall that Theorem F59(b) says that

$$\lim_{n\to\infty} \frac{(n!)^2\, 2^{2n}}{(2n)!\sqrt{n}} = \sqrt{\pi}.$$

Hence

$$\frac{L^2}{L} = \frac{1}{\sqrt{\pi}} \times \frac{e}{\sqrt{2}}, \quad \text{so} \quad L = \frac{e}{\sqrt{2\pi}}.$$

Thus we can rewrite statement (28) in the form

$$a_n = \frac{n^{n+(1/2)}}{e^{n-1}n!} \to \frac{e}{\sqrt{2\pi}} \quad \text{as } n \to \infty.$$

Hence, by the Combination Rules for sequences,

$$\frac{e^n n!}{n^n \sqrt{n}} \to \sqrt{2\pi} \quad \text{as } n \to \infty,$$

which can be rearranged to give Stirling's Formula:

$$n! \sim \sqrt{2\pi n}\,(n/e)^n \quad \text{as } n \to \infty.$$

■

Summary

In this unit you have studied the formal definition of what it means for a function f to be integrable on a closed interval $[a, b]$. You did this by studying lower and upper Riemann sums for f, which give lower and upper estimates for the area between the graph of f and the x-axis for $a \leq x \leq b$, by approximating the area with a sum of areas of rectangles. These rectangles are based on a collection of subintervals of $[a, b]$ known as a partition of $[a, b]$, and you saw that, by decreasing the length of these subintervals (that is, the mesh of the partition), we get increasingly accurate estimates for the area. If the supremum of the lower Riemann sums is equal to the infimum of the upper Riemann sums, then we say that the function f is integrable on $[a, b]$ and the common limit is the integral of f. You also saw that to prove that f is integrable it is sufficient to consider the sequence of standard partitions P_n where the interval is divided into n subintervals of equal lengths.

You then met the Fundamental Theorem of Calculus and saw that integration can be thought of as the process of finding a primitive, and hence as the opposite of differentiation. You studied a number of techniques for finding primitives, including integration by substitution, integration by parts and the use of a reduction formula.

You also met some useful inequalities that enable us to estimate various integrals when it is not possible to find an exact value. You saw how these can be used to prove Wallis' Formula, an approximation for π, and studied the Integral Test which gives a method based on integration for determining the convergence or divergence of certain series. Finally you met Stirling's Formula which gives an approximation for $n!$ and is proved using upper and lower Riemann sums.

Learning outcomes

After working through this unit, you should be able to:

- determine the *lower Riemann sum* $L(f, P)$ and the *upper Riemann sum* $U(f, P)$ for a given function f and *partition* P
- understand the definition of the *integral* $\int_a^b f$
- use upper and lower Riemann sums to determine whether a given function is *integrable*
- use basic rules for manipulating integrals
- state various sufficient conditions for a function to be integrable
- explain what is meant by a *primitive* of a function and understand the Fundamental Theorem of Calculus
- use the Fundamental Theorem of Calculus and the table of standard primitives to evaluate certain integrals
- use *integration by substitution* and *integration by parts*
- use the *reduction of order* method to evaluate certain integrals
- determine lower and upper estimates for given integrals
- state Wallis' Formula
- use the Integral Test to determine the convergence or divergence of certain series
- understand the connection between integration and Stirling's Formula for $n!$
- use Stirling's Formula to determine the behaviour of certain sequences involving factorials.

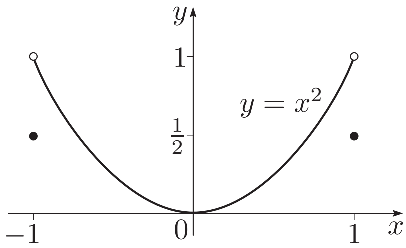
Table of standard primitives

$f(x)$	Primitive $F(x)$	Domain
$x^n,\ n \in \mathbb{Z} - \{-1\}$	$x^{n+1}/(n+1)$	$\mathbb{R}$ or $\mathbb{R} - \{0\}$
$x^\alpha,\ \alpha \neq -1$	$x^{\alpha+1}/(\alpha+1)$	$\mathbb{R}^+$
$a^x,\ a > 0$	$a^x/\log a$	$\mathbb{R}$
$\sin x$	$-\cos x$	$\mathbb{R}$
$\cos x$	$\sin x$	$\mathbb{R}$
$\tan x$	$\log(\sec x)$	$\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$
e^x	e^x	$\mathbb{R}$
$1/x$	$\log x$	$(0, \infty)$
$1/x$	$\log \lvert x \rvert$	$(-\infty, 0)$
$\log x$	$x \log x - x$	$(0, \infty)$
$\sinh x$	$\cosh x$	$\mathbb{R}$
$\cosh x$	$\sinh x$	$\mathbb{R}$
$\tanh x$	$\log(\cosh x)$	$\mathbb{R}$
$(a^2 - x^2)^{-1},\ a \neq 0$	$\dfrac{1}{2a} \log\left(\dfrac{a+x}{a-x}\right)$	$(-a, a)$
$(a^2 + x^2)^{-1},\ a \neq 0$	$\dfrac{1}{a} \tan^{-1}(x/a)$	$\mathbb{R}$
$(a^2 - x^2)^{-1/2},\ a \neq 0$	$\begin{cases} \sin^{-1}(x/a) \\ -\cos^{-1}(x/a) \end{cases}$	$(-a, a)$ $(-a, a)$
$(x^2 - a^2)^{-1/2},\ a \neq 0$	$\begin{cases} \log(x + (x^2 - a^2)^{1/2}) \\ \cosh^{-1}(x/a) \end{cases}$	(a, ∞) (a, ∞)
$(a^2 + x^2)^{-1/2},\ a \neq 0$	$\begin{cases} \log(x + (a^2 + x^2)^{1/2}) \\ \sinh^{-1}(x/a) \end{cases}$	$\mathbb{R}$ $\mathbb{R}$
$(a^2 - x^2)^{1/2},\ a \neq 0$	$\frac{1}{2}x(a^2 - x^2)^{1/2} + \frac{1}{2}a^2 \sin^{-1}(x/a)$	$(-a, a)$
$(x^2 - a^2)^{1/2},\ a \neq 0$	$\frac{1}{2}x(x^2 - a^2)^{1/2} - \frac{1}{2}a^2 \log(x + (x^2 - a^2)^{1/2})$	(a, ∞)
$(a^2 + x^2)^{1/2},\ a \neq 0$	$\frac{1}{2}x(a^2 + x^2)^{1/2} + \frac{1}{2}a^2 \log(x + (a^2 + x^2)^{1/2})$	$\mathbb{R}$
$e^{ax} \cos bx,\ a, b \neq 0$	$\dfrac{e^{ax}}{a^2 + b^2}(a \cos bx + b \sin bx)$	$\mathbb{R}$
$e^{ax} \sin bx,\ a, b \neq 0$	$\dfrac{e^{ax}}{a^2 + b^2}(a \sin bx - b \cos bx)$	$\mathbb{R}$

Solutions to exercises

Solution to Exercise F33

(a) The graph of f is shown below.



First, $\min f = 0$, since

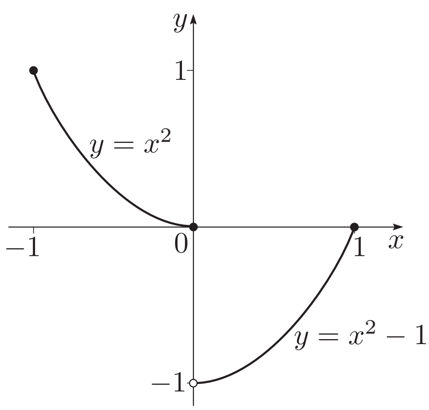
1. $f(x) \geq 0$, for all $x \in [-1, 1]$,
2. $f(0) = 0$.

Next, $\inf f = 0$, since f has minimum 0 on $[-1, 1]$.
Also, $\sup f = 1$, since

1. $f(x) \leq 1$, for all $x \in [-1, 1]$,
2. if $M' < 1$, then M' is not an upper bound for f on $[-1, 1]$ because the sequence $(1 - 1/n)$ is contained in $[-1, 1]$ and
$$f(1 - 1/n) = (1 - 1/n)^2 \to 1 \ \text{ as } n \to \infty,$$
so there exists $x' \in [-1, 1]$ such that $f(x') > M'$. (Alternatively, you could consider $f(-1 + 1/n)$.)

Finally, $\max f$ does not exist, since there is no point x such that $f(x) = 1$.

(b) The graph of f is shown below.



First, $\inf f = -1$, since

1. $f(x) \geq -1$, for all $x \in [-1, 1]$,
2. if $m' > -1$, then m' is not a lower bound for f on $[-1, 1]$ because the sequence $(1/n)$ is contained in $[-1, 1]$ and
$$f(1/n) = (1/n)^2 - 1 \to -1 \ \text{ as } n \to \infty,$$
so there exists $x' \in [-1, 1]$ such that $f(x') < m'$.

Next, $\min f$ does not exist, since there is no point x in $[-1, 1]$ such that $f(x) = -1$.
Also, $\max f = 1$, since

1. $f(x) \leq 1$, for all $x \in [-1, 1]$,
2. $f(-1) = 1$.

Finally, $\sup f = 1$, since f has maximum 1 on $[-1, 1]$.

Solution to Exercise F34

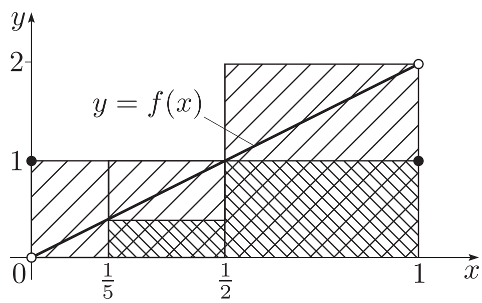
The interval $[-1, 2]$ has length 3; hence the 4 subintervals in this standard partition P of $[-1, 2]$ must each have length $\frac{3}{4}$. Thus the required standard partition of $[-1, 2]$ is

$$P = \left\{ \left[-1, -\tfrac{1}{4}\right], \left[-\tfrac{1}{4}, \tfrac{1}{2}\right], \left[\tfrac{1}{2}, \tfrac{5}{4}\right], \left[\tfrac{5}{4}, 2\right] \right\}.$$

The mesh of P is the common length of the subintervals, namely $\frac{3}{4}$.

Solution to Exercise F35

The function and the partition are illustrated below.



For the three subintervals in P, we have

$$\begin{aligned}
m_1 &= 0, & M_1 &= f(0) = 1, & \delta x_1 &= \tfrac{1}{5}, \\
m_2 &= f\left(\tfrac{1}{5}\right) = \tfrac{2}{5}, & M_2 &= f\left(\tfrac{1}{2}\right) = 1, & \delta x_2 &= \tfrac{3}{10}, \\
m_3 &= f\left(\tfrac{1}{2}\right) = 1, & M_3 &= 2, & \delta x_3 &= \tfrac{1}{2}.
\end{aligned}$$

(Note that m_3 is also equal to $f(1)$.)

Then

$$\begin{aligned} L(f,P) &= \sum_{i=1}^{3} m_i \delta x_i \\ &= \left(0 \times \frac{1}{5}\right) + \left(\frac{2}{5} \times \frac{3}{10}\right) + \left(1 \times \frac{1}{2}\right) \\ &= 0 + \frac{3}{25} + \frac{1}{2} \\ &= \frac{31}{50} \end{aligned}$$

and

$$\begin{aligned} U(f,P) &= \sum_{i=1}^{3} M_i \delta x_i \\ &= \left(1 \times \frac{1}{5}\right) + \left(1 \times \frac{3}{10}\right) + \left(2 \times \frac{1}{2}\right) \\ &= \frac{1}{5} + \frac{3}{10} + 1 \\ &= \frac{3}{2}. \end{aligned}$$

Solution to Exercise F36

Let

$$f(x) = x, \quad x \in [0,1],$$

and let P_n be the standard partition of $[0,1]$,

$$P_n = \left\{\left[0, \frac{1}{n}\right], \left[\frac{1}{n}, \frac{2}{n}\right], \dots, \left[\frac{n-1}{n}, 1\right]\right\}.$$

Since f is increasing we have, for $i = 1, 2, \dots, n$,

$$m_i = f\left(\frac{i-1}{n}\right) = \frac{i-1}{n},$$

$$M_i = f\left(\frac{i}{n}\right) = \frac{i}{n}$$

and

$$\delta x_i = \frac{1}{n}.$$

Thus

$$\begin{aligned} L(f,P_n) &= \sum_{i=1}^{n} m_i \delta x_i \\ &= \sum_{i=1}^{n} \left(\frac{i-1}{n} \times \frac{1}{n}\right) \\ &= \frac{1}{n^2}(1 + 2 + \dots + (n-1)) \\ &= \frac{1}{n^2} \times \frac{(n-1)n}{2} \\ &= \frac{1}{2}\left(1 - \frac{1}{n}\right), \end{aligned}$$

and also

$$\begin{aligned} U(f,P_n) &= \sum_{i=1}^{n} M_i \delta x_i \\ &= \sum_{i=1}^{n} \left(\frac{i}{n} \times \frac{1}{n}\right) \\ &= \frac{1}{n^2}(1 + 2 + \dots + n) \\ &= \frac{1}{n^2} \times \frac{n(n+1)}{2} \\ &= \frac{1}{2}\left(1 + \frac{1}{n}\right). \end{aligned}$$

(Here we have used the fact that the sum of the arithmetic series $1 + 2 + \dots + n$ is $\dfrac{n(n+1)}{2}$, for each $n \in \mathbb{N}$.)

Then, as $n \to \infty$, we have

$$\lim_{n\to\infty} L(f,P_n) = \frac{1}{2} \quad \text{and} \quad \lim_{n\to\infty} U(f,P_n) = \frac{1}{2}.$$

In particular,

$$\frac{1}{2} \leq \underline{\int_0^1} f \leq \overline{\int_0^1} f \leq \frac{1}{2}$$

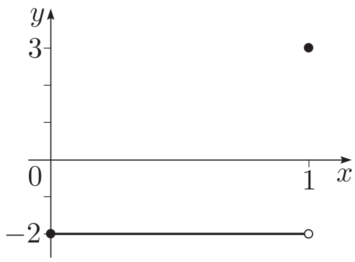
and hence

$$\underline{\int_0^1} f = \overline{\int_0^1} f = \frac{1}{2}.$$

Thus f is integrable on $[0,1]$ and $\displaystyle\int_0^1 f = \frac{1}{2}$.

Solution to Exercise F37

(a) The graph of f is shown below.



Let P_n be the standard partition of $[0,1]$ into n equal subintervals:

$$\left\{\left[0,\frac{1}{n}\right],\left[\frac{1}{n},\frac{2}{n}\right],\dots,\left[\frac{n-1}{n},1\right]\right\}.$$

For $i=1,2,\dots,n$, we have

$$m_i=\inf\{f(x):(i-1)/n\le x\le i/n\}=-2.$$

For $i=1,2,\dots,n-1$, we have

$$M_i=\sup\{f(x):(i-1)/n\le x\le i/n\}=-2.$$

Also, $M_n=3$ and

$$\delta x_i=\frac{1}{n},\quad \text{for } i=1,2,\dots,n.$$

Hence

$$\begin{aligned}L(f,P_n)&=\sum_{i=1}^{n}m_i\,\delta x_i\\&=\sum_{i=1}^{n}\left(-2\times\frac{1}{n}\right)\\&=n\times\left(\frac{-2}{n}\right)=-2,\end{aligned}$$

$$\begin{aligned}U(f,P_n)&=\sum_{i=1}^{n}M_i\,\delta x_i\\&=\sum_{i=1}^{n-1}M_i\,\delta x_i+M_n\,\delta x_n\\&=\sum_{i=1}^{n-1}\left(-2\times\frac{1}{n}\right)+\left(3\times\frac{1}{n}\right)\\&=(n-1)\left(\frac{-2}{n}\right)+\frac{3}{n}\\&=-2+\frac{5}{n}\to-2\ \text{ as } n\to\infty.\end{aligned}$$

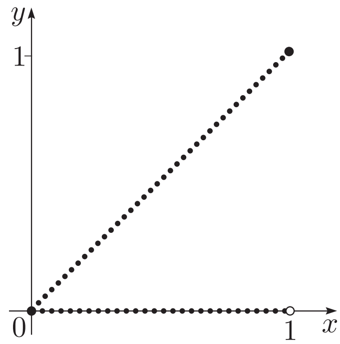
Since $\|P_n\|\to 0$ as $n\to\infty$, and

$$\lim_{n\to\infty}L(f,P_n)=\lim_{n\to\infty}U(f,P_n)=-2,$$

it follows from Theorem F45 that f is integrable on $[0,1]$ and

$$\int_0^1 f=-2.$$

(b) The graph of f is shown below.



Let P_n be the standard partition of $[0,1]$ into n equal subintervals:

$$\left\{\left[0,\frac{1}{n}\right],\left[\frac{1}{n},\frac{2}{n}\right],\dots,\left[\frac{n-1}{n},1\right]\right\}.$$

Each subinterval

$$\left[\frac{i-1}{n},\frac{i}{n}\right],\quad \text{for } i=1,2,\dots,n,$$

contains both rational and irrational points:

at the rational points, $f(x)=x$,
at the irrational points, $f(x)=0$.

Hence, for $i=1,2,\dots,n$,

$$m_i=0\quad\text{and}\quad M_i=\frac{i}{n}.$$

Also,

$$\delta x_i=\frac{1}{n},\quad \text{for } i=1,2,\dots,n.$$

Hence

$$\begin{aligned}L(f,P_n)&=\sum_{i=1}^{n} m_i\,\delta x_i\\&=\sum_{i=1}^{n}\left(0\times\frac{1}{n}\right)=0,\\U(f,P_n)&=\sum_{i=1}^{n} M_i\,\delta x_i\\&=\sum_{i=1}^{n}\left(\frac{i}{n}\times\frac{1}{n}\right)\\&=\frac{1}{n^2}\sum_{i=1}^{n} i\\&=\frac{1}{n^2}\times\frac{n(n+1)}{2}\\&=\frac{1}{2}+\frac{1}{2n}\to\frac{1}{2}\quad\text{as } n\to\infty.\end{aligned}$$

Since $\|P_n\|\to 0$ as $n\to\infty$, but

$$\lim_{n\to\infty} L(f,P_n)=0\neq\tfrac{1}{2}=\lim_{n\to\infty} U(f,P_n),$$

it follows from Theorem F46 that f is not integrable on $[0,1]$.

Solution to Exercise F38

(a) We have

$$\begin{aligned}F'(x)&=\frac{1}{x+(x^2-4)^{1/2}}\left(1+\tfrac{1}{2}(x^2-4)^{-1/2}\,2x\right)\\&=\frac{(x^2-4)^{-1/2}\left((x^2-4)^{1/2}+x\right)}{x+(x^2-4)^{1/2}}\\&=(x^2-4)^{-1/2}\\&=f(x),\end{aligned}$$

as required.

(b) We have

$$\begin{aligned}F'(x)&=\frac{1}{1+\sinh^2 x}\cosh x\\&=\frac{\cosh x}{\cosh^2 x}\\&=\operatorname{sech} x\\&=f(x),\end{aligned}$$

as required.

Solution to Exercise F39

(a) From the Fundamental Theorem of Calculus and the table of standard primitives, we deduce that

$$\begin{aligned}&\int_0^4 (x^2+9)^{1/2}\,dx\\&=\left[\tfrac{1}{2}x(x^2+9)^{1/2}+\tfrac{9}{2}\log(x+(x^2+9)^{1/2})\right]_0^4\\&=10+\tfrac{9}{2}\log 9-\tfrac{9}{2}\log 3\\&=10+\tfrac{9}{2}\log 3.\end{aligned}$$

(b) From the Fundamental Theorem of Calculus and the table of standard primitives, we deduce that

$$\begin{aligned}\int_1^e \log x\,dx&=\left[x\log x-x\right]_1^e\\&=(e-e)-(0-1)=1.\end{aligned}$$

Solution to Exercise F40

Using the table of standard primitives and the Combination Rules, we obtain the following primitives.

(a) $$\begin{aligned}F(x)&=4(x\log x-x)-2\left(\tfrac{1}{2}\tan^{-1}(x/2)\right)\\&=4(x\log x-x)-\tan^{-1}(x/2)\end{aligned}$$

(b) $$\begin{aligned}F(x)&=2\left(\tfrac{1}{3}\log(\sec 3x)\right)+\frac{e^{2x}}{2^2+1^2}(2\cos x+\sin x)\\&=\tfrac{2}{3}\log(\sec 3x)+\tfrac{1}{5}e^{2x}(2\cos x+\sin x)\end{aligned}$$

(These results can be checked by differentiation.)

Solution to Exercise F41

(a) Take $u=\sin 3x$; then

$$\frac{du}{dx}=3\cos 3x,\quad\text{so}\quad du=3\cos 3x\,dx.$$

Hence

$$\begin{aligned}\int\sin(\sin 3x)\cos 3x\,dx&=\tfrac{1}{3}\int\sin u\,du\\&=-\tfrac{1}{3}\cos u\\&=-\tfrac{1}{3}\cos(\sin 3x).\end{aligned}$$

(b) Taking $u=2+3x^3$, we obtain

$$\int x^2(2+3x^3)^7\,dx=\tfrac{1}{72}(2+3x^3)^8.$$

(c) Taking $u = 2x^2$, we obtain

$$\int x\sin(2x^2)\,dx = -\tfrac{1}{4}\cos(2x^2).$$

(d) Using equation (17), we obtain

$$\int x/(2+3x^2)\,dx = \tfrac{1}{6}\log(2+3x^2).$$

Solution to Exercise F42

Let $u = 1 + e^x$; then

$$\frac{du}{dx} = e^x, \quad \text{so} \quad du = e^x\,dx.$$

Also,

$$\begin{aligned} &\text{when } x = 0, \quad u = 2,\\ &\text{when } x = 1, \quad u = 1 + e. \end{aligned}$$

Hence

$$\begin{aligned} \int_0^1 \frac{e^x}{(1+e^x)^2}\,dx &= \int_2^{1+e} \frac{du}{u^2}\\ &= \left[\frac{-1}{u}\right]_2^{1+e}\\ &= -\frac{1}{1+e} + \frac{1}{2} = \frac{e-1}{2(1+e)}. \end{aligned}$$

Solution to Exercise F43

(a) Let $u = (x-1)^{1/2}$, so $x = u^2 + 1$; then

$$\frac{dx}{du} = 2u, \quad \text{so} \quad dx = 2u\,du.$$

Hence

$$\begin{aligned} &\int \frac{dx}{3(x-1)^{3/2} + x(x-1)^{1/2}}\\ &= \int \frac{2u}{3u^3 + (u^2+1)u}\,du\\ &= \int \frac{2}{4u^2+1}\,du\\ &= 2\int \frac{du}{(2u)^2+1}\\ &= \tan^{-1}(2u)\\ &= \tan^{-1}(2(x-1)^{1/2}). \end{aligned}$$

(b) Let $u = \sqrt{1+e^x}$, so $x = \log(u^2-1)$; then

$$\frac{dx}{du} = \frac{2u}{u^2-1}, \quad \text{so} \quad dx = \frac{2u}{u^2-1}\,du.$$

Also,

$$\begin{aligned} &\text{when } x = 0, \quad u = \sqrt{2},\\ &\text{when } x = \log 3, \quad u = 2. \end{aligned}$$

Hence

$$\begin{aligned} \int_0^{\log 3} e^x\sqrt{1+e^x}\,dx &= \int_{\sqrt{2}}^2 (u^2-1)u\frac{2u}{u^2-1}\,du\\ &= \int_{\sqrt{2}}^2 2u^2\,du\\ &= \left[\tfrac{2}{3}u^3\right]_{\sqrt{2}}^2\\ &= (16 - 4\sqrt{2})/3. \end{aligned}$$

Solution to Exercise F44

(a) Here we use integration by parts, with

$$f(x) = \log x \quad \text{and} \quad g'(x) = x^{1/3};$$

then

$$f'(x) = 1/x \quad \text{and} \quad g(x) = \tfrac{3}{4}x^{4/3}.$$

Hence

$$\begin{aligned} \int x^{1/3}\log x\,dx &= \tfrac{3}{4}x^{4/3}\log x - \tfrac{3}{4}\int x^{4/3}x^{-1}\,dx\\ &= \tfrac{3}{4}x^{4/3}\log x - \tfrac{3}{4}\int x^{1/3}\,dx\\ &= \tfrac{3}{4}x^{4/3}\log x - \tfrac{9}{16}x^{4/3}. \end{aligned}$$

(b) We use integration by parts twice. On each occasion we differentiate the power function and integrate the trigonometric function.

We have

$$\begin{aligned} \int_0^{\pi/2} x^2\cos x\,dx &= \left[x^2\sin x\right]_0^{\pi/2}\\ &\qquad - \int_0^{\pi/2} 2x\sin x\,dx\\ &= \frac{\pi^2}{4} - 2\int_0^{\pi/2} x\sin x\,dx \end{aligned}$$

and

$$\begin{aligned} \int_0^{\pi/2} x\sin x\,dx &= \left[x(-\cos x)\right]_0^{\pi/2}\\ &\qquad - \int_0^{\pi/2}(-\cos x)\,dx\\ &= 0 + \left[\sin x\right]_0^{\pi/2} = 1. \end{aligned}$$

It follows that

$$\int_0^{\pi/2} x^2 \cos x \, dx = \frac{\pi^2}{4} - 2.$$

Solution to Exercise F45

(a) $I_0 = \displaystyle\int_0^1 e^x \, dx = [e^x]_0^1 = e - 1.$

(b) Using integration by parts, we obtain

$$\begin{aligned} I_n &= \int_0^1 e^x x^n \, dx \\ &= \Big[e^x x^n\Big]_0^1 - \int_0^1 e^x n x^{n-1} \, dx \\ &= e - nI_{n-1}, \quad \text{for } n \geq 1. \end{aligned}$$

(c) Using the solution to part (b) with $n = 1, 2, 3, 4$ in turn, we obtain

$$\begin{aligned} I_1 &= e - I_0 = e - (e-1) = 1, \\ I_2 &= e - 2I_1 = e - 2, \\ I_3 &= e - 3I_2 = e - 3(e-2) = 6 - 2e, \\ I_4 &= e - 4I_3 = e - 4(6-2e) = 9e - 24. \end{aligned}$$

Solution to Exercise F46

(a) Since $\sin(1/x^{10}) \leq 1$, we have

$$x \sin(1/x^{10}) \leq x, \quad \text{for } x \in [1,3].$$

Thus it follows from Inequality Rule (a) that

$$\begin{aligned} \int_1^3 x \sin(1/x^{10}) \, dx &\leq \int_1^3 x \, dx \\ &= \left[\tfrac{1}{2}x^2\right]_1^3 \\ &= \tfrac{1}{2}(9-1) \\ &= 4. \end{aligned}$$

(b) If $x \in \left[0, \frac{1}{2}\right]$, then

$$1 = e^0 \leq e^{x^2} \leq e^{(1/2)^2} = e^{1/4},$$

because the function $x \longmapsto e^{x^2}$ is increasing on $\left[0, \frac{1}{2}\right]$.

Thus it follows from Inequality Rule (b) that

$$\tfrac{1}{2} \leq \int_0^{1/2} e^{x^2} \, dx \leq \tfrac{1}{2} e^{1/4}.$$

Solution to Exercise F47

(a) Since

$$|\sin(1/x)| \leq 1, \quad \text{for } x \in [1,4],$$

and

$$2 + \cos(1/x) \geq 1, \quad \text{for } x \in [1,4],$$

it follows that

$$\left|\frac{\sin(1/x)}{2 + \cos(1/x)}\right| \leq 1, \quad \text{for } x \in [1,4].$$

Hence, by the Triangle Inequality for integrals,

$$\left|\int_1^4 \frac{\sin(1/x)}{2 + \cos(1/x)} \, dx\right| \leq \int_1^4 dx = 4 - 1 = 3.$$

(b) Since

$$\tan x \geq 0, \quad \text{for } x \in [0, \pi/4],$$

and

$$3 - \sin(x^2) \geq 2, \quad \text{for } x \in [0, \pi/4],$$

it follows that

$$0 \leq \frac{\tan x}{3 - \sin(x^2)} \leq \tfrac{1}{2} \tan x, \quad \text{for } x \in [0, \pi/4].$$

Hence, by the Triangle Inequality for integrals and Inequality Rule (a),

$$\begin{aligned} \left|\int_0^{\pi/4} \frac{\tan x}{3 - \sin(x^2)} \, dx\right| &\leq \int_0^{\pi/4} \left|\frac{\tan x}{3 - \sin(x^2)}\right| dx \\ &\leq \int_0^{\pi/4} \tfrac{1}{2} \tan x \, dx \\ &= \left[\tfrac{1}{2} \log(\sec x)\right]_0^{\pi/4} \\ &= \tfrac{1}{2}(\log(\sec \pi/4) - \log 1) \\ &= \tfrac{1}{2} \log(\sqrt{2}) \\ &= \tfrac{1}{4} \log 2. \end{aligned}$$

Solution to Exercise F48

(a) We have

$$a_1 = \frac{2}{1}\cdot\frac{2}{3} = \frac{4}{3},$$
$$a_2 = \frac{2}{1}\cdot\frac{2}{3}\cdot\frac{4}{3}\cdot\frac{4}{5} = \frac{64}{45},$$
$$a_3 = \frac{2}{1}\cdot\frac{2}{3}\cdot\frac{4}{3}\cdot\frac{4}{5}\cdot\frac{6}{5}\cdot\frac{6}{7} = \frac{256}{175},$$
$$b_1 = \frac{(1!)^2\, 2^2}{2!\sqrt{1}} = 2,$$
$$b_2 = \frac{(2!)^2\, 2^4}{4!\sqrt{2}} = \frac{4}{3}\sqrt{2},$$
$$b_3 = \frac{(3!)^2\, 2^6}{6!\sqrt{3}} = \frac{16}{15}\sqrt{3}.$$

(b) Using the results of part (a),

$$b_1^2 = 4 = 3a_1,$$
$$b_2^2 = \frac{32}{9} = \frac{5}{2}a_2,$$
$$b_3^2 = \frac{256}{75} = \frac{7}{3}a_3,$$

as required.

(c) We have

$$b_n^2 = \frac{(n!)^4\, 2^{4n}}{((2n)!)^2\, n}.$$

We now try to express

$$a_n = \frac{2\cdot 2\cdot 4\cdot 4\cdot \cdots \cdot (2n)(2n)}{1\cdot 3\cdot 3\cdot 5\cdot \cdots \cdot (2n-1)(2n+1)}$$

in terms of factorials.

The numerator is a product of $2n$ even numbers. Taking a factor 2 from each term, we deduce that

$$\begin{aligned}&2\cdot 2\cdot 4\cdot 4\cdot \cdots \cdot (2n)(2n)\\ &\quad = 2^{2n}(1\cdot 1\cdot 2\cdot 2\cdot \cdots \cdot n\cdot n) = 2^{2n}(n!)^2.\end{aligned}$$

The denominator of a_n cannot be treated in quite the same way, as all its factors are odd. To relate it to factorials, we introduce the missing even factors:

$$\begin{aligned}&1\cdot 3\cdot 3\cdot 5\cdot 5\cdot \cdots \cdot (2n-1)(2n+1)\\ &= \frac{1\cdot 2\cdot 2\cdot 3\cdot 3\cdot 4\cdot 4\cdot \cdots \cdot (2n-1)(2n)(2n)(2n+1)}{2\cdot 2\cdot 4\cdot 4\cdot \cdots \cdot (2n)(2n)}\\ &= \frac{((2n)!)^2\,(2n+1)}{2^{2n}(n!)^2}.\end{aligned}$$

It follows that

$$\begin{aligned}a_n &= 2^{2n}(n!)^2 \Big/ \frac{((2n)!)^2\,(2n+1)}{2^{2n}(n!)^2}\\ &= \frac{2^{4n}(n!)^4}{((2n)!)^2(2n+1)}\\ &= b_n^2\left(\frac{n}{2n+1}\right).\end{aligned}$$

Hence

$$b_n^2 = \left(\frac{2n+1}{n}\right)a_n,$$

as required.

Solution to Exercise F49

Let $u = \log x$; then

$$\frac{du}{dx} = \frac{1}{x}, \quad \text{so} \quad du = \frac{dx}{x}.$$

Hence

$$\int \frac{dx}{x(\log x)^2} = \int \frac{du}{u^2} = -\frac{1}{u} = -\frac{1}{\log x}.$$

Now let

$$f(x) = \frac{1}{x(\log x)^2} \quad (x \in [2,\infty)).$$

Then f is positive and decreasing on $[2,\infty)$, and

$$f(x) \to 0 \text{ as } x \to \infty.$$

Also, for $n \geq 2$,

$$\begin{aligned}\int_2^n f &= \int_2^n \frac{dx}{x(\log x)^2}\\ &= \left[-\frac{1}{\log x}\right]_2^n\\ &= \frac{1}{\log 2} - \frac{1}{\log n} \leq \frac{1}{\log 2}.\end{aligned}$$

Since the sequence of integrals $\left(\int_2^n f\right)_2^\infty$ is bounded above, it follows from part (a) of the Integral Test that the series $\displaystyle\sum_{n=2}^\infty \frac{1}{n(\log n)^2}$ converges.

Solution to Exercise F50

The values are as follows.

n	$n!$
30	$2.652\ldots \times 10^{32}$
40	$8.159\ldots \times 10^{47}$
50	$3.041\ldots \times 10^{64}$
60	$8.320\ldots \times 10^{81}$

Solution to Exercise F51

The values are as follows.

n	$\sqrt{2\pi n}\,(n/e)^n$
5	$118.019\ldots$
10	$3.598\ldots \times 10^{6}$
50	$3.036\ldots \times 10^{64}$

Solution to Exercise F52

In each part we approximate the factorials using Stirling's Formula.

(a)
$$\begin{aligned}\binom{300}{150}\frac{1}{2^{300}} &= \frac{300!}{150!\,150!\,2^{300}}\\ &\approx \frac{\sqrt{600\pi}\,(300/e)^{300}}{\left(\sqrt{300\pi}\,(150/e)^{150}\right)^2 2^{300}}\\ &= \frac{\sqrt{600\pi}}{300\pi}\\ &= \frac{\sqrt{6}}{30\sqrt{\pi}}\\ &= 0.046 \quad \text{(to 2 s.f.)}.\end{aligned}$$

(b)
$$\begin{aligned}\frac{300!}{(100!)^3}\frac{1}{3^{300}} &\approx \frac{\sqrt{600\pi}\,(300/e)^{300}}{(200\pi)^{3/2}\,(100/e)^{300}\,3^{300}}\\ &= \frac{\sqrt{600\pi}}{(200\pi)^{3/2}}\\ &= \frac{\sqrt{3}}{200\pi}\\ &= 0.0028 \quad \text{(to 2 s.f.)}.\end{aligned}$$

Solution to Exercise F53

We have
$$\binom{4n}{2n} = \frac{(4n)!}{((2n)!)^2} \quad \text{and} \quad \binom{2n}{n} = \frac{(2n)!}{(n!)^2},$$
so
$$\binom{4n}{2n}\bigg/\binom{2n}{n} = \frac{(4n)!(n!)^2}{((2n)!)^3}.$$
By Stirling's Formula, and the Product and Quotient Rules for $\sim$, we obtain
$$\begin{aligned}\binom{4n}{2n}\bigg/\binom{2n}{n} &\sim \frac{\sqrt{8\pi n}\,(4n/e)^{4n}\left(\sqrt{2\pi n}\,(n/e)^n\right)^2}{\left(\sqrt{4\pi n}\,(2n/e)^{2n}\right)^3}\\ &= \frac{\sqrt{8\pi n}\,4^{4n}\,2\pi n}{(4\pi n)^{3/2}\,2^{6n}}\\ &= \frac{2\sqrt{8}}{8}\frac{4^{4n}}{2^{6n}}\\ &= \frac{1}{\sqrt{2}}\,2^{2n}.\end{aligned}$$
Hence $\lambda = 1/\sqrt{2}$.

Solution to Exercise F54

Using Stirling's Formula, we obtain
$$\frac{n^n}{n!} \sim \frac{n^n}{\sqrt{2\pi n}\,(n/e)^n} = \frac{e^n}{\sqrt{2\pi n}}.$$
Thus, by the hint,
$$\left(\frac{n^n}{n!}\right)^{1/n} \sim \left(\frac{e^n}{\sqrt{2\pi n}}\right)^{1/n} = \frac{e}{\left(\sqrt{2\pi}\right)^{1/n}\sqrt{n^{1/n}}}.$$
We know (from Worked Exercise D26 and Exercise D34 in Subsection 3.3 of Unit D2) that, for any positive number a,
$$a^{1/n} \to 1 \quad \text{as } n \to \infty$$
and
$$n^{1/n} \to 1 \quad \text{as } n \to \infty.$$
Hence
$$\left(\sqrt{2\pi}\right)^{1/n} \to 1 \quad \text{as } n \to \infty$$
and
$$\sqrt{n^{1/n}} \to 1 \quad \text{as } n \to \infty.$$

It follows that

$$\left(\frac{n^n}{n!}\right)^{1/n} \sim e \quad \text{as } n \to \infty;$$

that is,

$$\left(\frac{n^n}{n!}\right)^{1/n} \to e \quad \text{as } n \to \infty.$$

Unit F4

Power series

Introduction

The evaluation of functions is of great importance. If you are dealing with a *polynomial* function, then the calculation of its values is just a matter of arithmetic. For example, if

$$f(x) = 1 + \tfrac{1}{2}x - \tfrac{1}{2}x^2 - \tfrac{1}{6}x^3 + \tfrac{1}{4}x^4,$$

then

$$f(1) = 1 + \tfrac{1}{2} - \tfrac{1}{2} - \tfrac{1}{6} + \tfrac{1}{4} = \tfrac{13}{12}.$$

On the other hand, the sine function is different. There is no way of calculating most of its values exactly just by the use of arithmetic, but it is important to be able to estimate them accurately because they arise in the solutions of many practical problems.

This unit is concerned with a procedure for calculating approximate values of functions, like the sine function, which cannot be found exactly. In studying this procedure, you will use many of the ideas and results you met in earlier analysis units, especially those relating to sequences and series.

You will see that a certain sequence of polynomials, known as *Taylor polynomials*, can be used to calculate approximate values of functions to any desired degree of accuracy, and that many functions can be represented as a sum of a convergent series of powers of x, known as a *Taylor series*. For example, the polynomial $p(x) = x - x^3/6$ approximates $f(x) = \sin x$ to within 5×10^{-6} for all x in the interval $[0, 0.1]$, and we can represent the sine function by the following convergent series:

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots, \quad \text{for } x \in \mathbb{R}.$$

As an application of these ideas, we show later in the unit how a method based on Taylor series can be used to estimate π to any desired number of decimal places.

1 Taylor polynomials

In this section you will meet the definition of the Taylor polynomial $T_n(x)$ at a point a of a function f, and study several particular functions for which Taylor polynomials appear to provide good approximations.

1.1 What are Taylor polynomials?

Let f be a function defined on an open interval I. Throughout this unit, we assume that a is a particular point of I and seek polynomial functions which provide good approximations to f near the point a.

If f is continuous at a, then the value $f(a)$ is an approximation to the value of $f(x)$ when x is near a, by the definition of continuity; that is,

$$f(x) \approx f(a), \quad \text{for } x \text{ near } a.$$

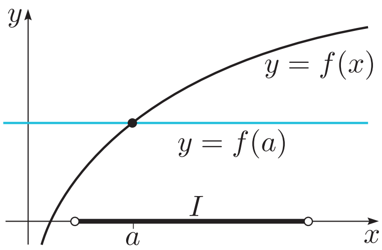


Figure 1 Approximating $f(x)$ by $f(a)$

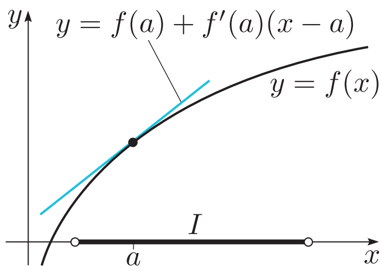


Figure 2 The tangent approximation at a to f

In geometric terms, this means that we can approximate the graph $y = f(x)$ near a by the horizontal line $y = f(a)$ through the point $(a, f(a))$, as shown in Figure 1. Usually this does not give a very good approximation.

However, if the function f is differentiable on I, then we can obtain what is usually a better approximation by using the tangent at $(a, f(a))$ instead of the horizontal line, as shown in Figure 2. We can think of the tangent at $(a, f(a))$ as the *line of best approximation* to the graph near a. The tangent to the graph at $(a, f(a))$ has equation

$$\frac{y - f(a)}{x - a} = f'(a), \quad \text{that is,} \quad y = f(a) + f'(a)(x - a).$$

So, for x near a, we can write

$$f(x) \approx f(a) + f'(a)(x - a).$$

This approximation is called the **tangent approximation** at a to f.

Note that the function f and the approximating linear function

$$x \longmapsto f(a) + f'(a)(x - a)$$

have the same value at a and the same first derivative at a (that is, their graphs have the same gradient at a), so this gives a better approximation to f near a than the constant function $f(a)$ when the gradient of the graph is non-zero.

Worked Exercise F40

Determine the tangent approximation to the function $f(x) = e^x$ at the point 0.

Solution

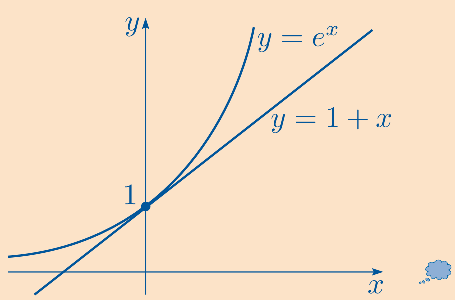
Here $a = 0$ and

$$f(x) = e^x, \quad f(0) = 1$$
$$f'(x) = e^x, \quad f'(0) = 1.$$

Hence the tangent approximation to f at 0 is

$$f(x) \approx f(0) + f'(0)(x - 0) = 1 + x.$$

This is illustrated in the graph below.



Exercise F55

Determine the tangent approximation to each of the following functions f at the given point a.

(a) $f(x) = e^x, \quad a = 2.$ (b) $f(x) = \cos x, \quad a = 0.$

So far we have seen two approximations to $f(x)$ for x near a:

$$f(x) \approx f(a) \quad \text{(a constant function)},$$
$$f(x) \approx f(a) + f'(a)(x-a) \quad \text{(a linear function)}.$$

If the function f is twice differentiable on I, then we can consider the quadratic function

$$p(x) = f(a) + f'(a)(x-a) + \tfrac{1}{2}f''(a)(x-a)^2,$$

which is chosen to satisfy $p(a) = f(a)$, $p'(a) = f'(a)$ and $p''(a) = f''(a)$, as you can check by differentiating p to give

$$p'(x) = f'(a) + f''(a)(x-a), \quad \text{and} \quad p''(x) = f''(a).$$

It is plausible that, for x near a,

$$f(x) \approx f(a) + f'(a)(x-a) + \tfrac{1}{2}f''(a)(x-a)^2$$

will usually be a better approximation to f near a than the constant or linear approximations. More generally, if the function f is n-times differentiable on I, then we can find a polynomial of degree n whose value at a and first n derivatives at a are equal to those of f, and this polynomial will usually provide an even better approximation.

Definition

Let f be n-times differentiable on an open interval containing the point a. Then the **Taylor polynomial of degree n at a for f** is the polynomial

$$T_n(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n,$$

that is,

$$T_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^n.$$

Remarks

1. Of course, the Taylor polynomial depends on the point a and the function f as well as on n and x, but for brevity our notation $T_n(x)$ does not reflect this. Usually a and f will be clear from the context.
2. The coefficients in the definition of T_n have been chosen so that
$$T_n(a) = f(a), \quad T_n'(a) = f'(a), \quad \ldots, \quad T_n^{(n)}(a) = f^{(n)}(a).$$
Thus the functions f and T_n have the same value at a and have equal derivatives at a for all orders up to and including n, in the same way as for the linear and quadratic functions above. Indeed, $T_0(x)$, $T_1(x)$ and $T_2(x)$ are respectively the constant, linear and quadratic approximations at a to f discussed previously.
3. It follows from the definition that Taylor polynomials for successive values of n satisfy the recurrence relation
$$T_n(x) = T_{n-1}(x) + \frac{f^{(n)}(a)}{n!}(x-a)^n, \quad n \geq 1.$$
4. A Taylor polynomial has degree n if it is based on derivatives of f up to order n. This use of the word 'degree' differs from the usual definition of the degree of a polynomial, which is the largest exponent in the polynomial.
5. Some texts refer to a Taylor polynomial *about* a instead of *at* a.

Worked Exercise F41

Determine the Taylor polynomials $T_1(x)$, $T_2(x)$ and $T_3(x)$ at the following points a for the function $f(x) = \sin x$.

(a) $a = 0$ (b) $a = \pi/2$

Solution

To find the Taylor polynomial of degree n at a, we need to work out the first n derivatives of f at a.

We have

$$\begin{aligned}
f(x) &= \sin x, & f(0) &= 0, & f(\pi/2) &= 1\\
f'(x) &= \cos x, & f'(0) &= 1, & f'(\pi/2) &= 0\\
f''(x) &= -\sin x, & f''(0) &= 0, & f''(\pi/2) &= -1\\
f^{(3)}(x) &= -\cos x, & f^{(3)}(0) &= -1, & f^{(3)}(\pi/2) &= 0.
\end{aligned}$$

It is convenient to use the recurrence relation for $T_n(x)$ when determining Taylor polynomials for successive values of n.

(a) Hence, at $a = 0$,

$$T_1(x) = f(0) + f'(0)x = x$$
$$T_2(x) = T_1(x) + \frac{f''(0)}{2!}x^2 = x$$
$$T_3(x) = T_2(x) + \frac{f^{(3)}(0)}{3!}x^3 = x - \tfrac{1}{6}x^3.$$

Note that $T_2(x)$, the Taylor polynomial of degree 2 at 0, is a polynomial of degree 1 because $f''(0) = 0$.

(b) At $a = \pi/2$ we have

$$T_1(x) = f(\pi/2) + f'(\pi/2)(x - \pi/2) = 1$$
$$\begin{aligned} T_2(x) &= T_1(x) + \frac{f''(\pi/2)}{2!}(x - \pi/2)^2 \\ &= 1 - \tfrac{1}{2}(x - \pi/2)^2 \end{aligned}$$
$$\begin{aligned} T_3(x) &= T_2(x) + \frac{f^{(3)}(\pi/2)}{3!}(x - \pi/2)^3 \\ &= 1 - \tfrac{1}{2}(x - \pi/2)^2. \end{aligned}$$

We do not usually multiply out brackets in such Taylor polynomials, since that would make the results less clear.

Exercise F56

Determine the Taylor polynomials $T_1(x)$, $T_2(x)$ and $T_3(x)$ for each of the following functions f at the given point a.

(a) $f(x) = e^x$, $a = 2$. (b) $f(x) = \cos x$, $a = 0$.

Exercise F57

Determine the Taylor polynomial of degree 4 for each of the following functions f at the given point a.

(a) $f(x) = \log(1 + x)$, $a = 0$. (b) $f(x) = \sin x$, $a = \pi/4$.

(c) $f(x) = 1 + \frac{1}{2}x - \frac{1}{2}x^2 - \frac{1}{6}x^3 + \frac{1}{4}x^4$, $a = 0$.

Exercise F58

Let T_3 be the Taylor polynomial of degree 3 at 0 for $f(x) = \sin x$, as calculated in Worked Exercise F41(a). Use a calculator to show that

$$|\sin(0.1) - T_3(0.1)| < 1 \times 10^{-7}.$$

(Remember to set the calculator to use angles in radians.)

1.2 Approximation by Taylor polynomials

We now look at two specific functions in order to investigate the assertion that Taylor polynomials provide good approximations for a large class of functions.

The function $f(x) = \sin x$

In Worked Exercise F41(a) we found the Taylor polynomials $T_1(x)$, $T_2(x)$ and $T_3(x)$ for the function $f(x) = \sin x$ at the point $a = 0$. By calculating higher derivatives of f at 0, we can show that the Taylor polynomials of degrees $1, 2, \ldots, 8$ at 0 for f are as follows.

$$T_1(x) = T_2(x) = x, \qquad T_3(x) = T_4(x) = x - \frac{x^3}{3!},$$

$$T_5(x) = T_6(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!}, \quad T_7(x) = T_8(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!}.$$

The graphs in Figure 3 illustrate how the approximation to $f(x)$ given by $T_n(x)$ improves as n increases.

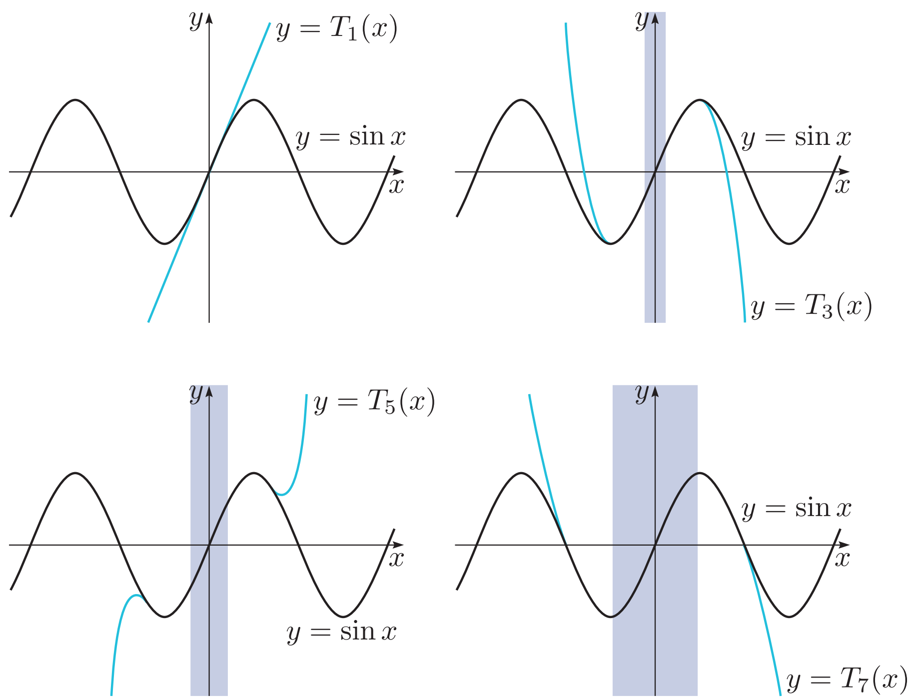


Figure 3 The graphs of Taylor polynomials at 0 for $f(x) = \sin x$

For example, the graph of T_5 appears to be very close to the graph of the sine function over the interval $(-\pi/2, \pi/2)$, so $T_5(x)$ seems to be a good approximation to $\sin x$ in this interval.

It also appears that, as the degree of the Taylor polynomial increases, the interval over which its graph is a good approximation to that of the sine function becomes longer. For instance, in the graphs in Figure 3, the shaded area covers the interval of the x-axis on which the Taylor polynomial $T_n(x)$ agrees with $\sin x$ to three decimal places.

Worked Exercise F42

Determine the Taylor polynomial of degree n at 0 for the function $f(x) = \sin x$.

Solution

We have

$$\begin{aligned}
f(x) &= \sin x, & f(0) &= 0\\
f'(x) &= \cos x, & f'(0) &= 1\\
f''(x) &= -\sin x, & f''(0) &= 0\\
f^{(3)}(x) &= -\cos x, & f^{(3)}(0) &= -1\\
f^{(4)}(x) &= \sin x, & f^{(4)}(0) &= 0
\end{aligned}$$

and in general, for $k = 0, 1, 2, \ldots,$

$$f^{(2k)}(0) = 0 \quad \text{and} \quad f^{(2k+1)}(0) = (-1)^k.$$

Hence, for $m = 0, 1, 2, \ldots,$

$$\begin{aligned}
T_{2m+1}(x) &= \sum_{k=0}^{m} (-1)^k \frac{x^{2k+1}}{(2k+1)!}\\
&= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + (-1)^m \frac{x^{2m+1}}{(2m+1)!}
\end{aligned}$$

and $T_{2m+2}(x) = T_{2m+1}(x)$.

So if n takes either of the values $2m+1$ or $2m+2$, then we have

$$T_n(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + (-1)^m \frac{x^{2m+1}}{(2m+1)!}.$$

The function $f(x) = 1/(1-x)$

By repeated differentiation of $f(x) = 1/(1-x)$, we can verify that

$$f^{(k)}(x) = \frac{k!}{(1-x)^{k+1}}, \quad \text{for } k \in \mathbb{N}.$$

Thus in particular $f^{(k)}(0) = k!$, so the Taylor polynomial of degree n at 0 for f is

$$\begin{aligned} T_n(x) &= \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k \\ &= \sum_{k=0}^{n} x^k = 1 + x + x^2 + \cdots + x^n. \end{aligned}$$

Figure 4 shows the graphs of the Taylor polynomials of degrees 1, 2, 4 and 7 at 0 for f.

The graphs show that the nature of the approximation is very different from that in the previous example. For the sine function, the interval over which the approximation is good seems to expand indefinitely as the degree of the polynomial increases. For $f(x) = 1/(1-x)$, however, the interval of good approximation always seems to be contained in the interval $(-1, 1)$.

For this function f, the Taylor polynomials $T_n(x)$ at 0 are the nth partial sums of the geometric series $\sum_{n=0}^{\infty} x^n$. This series converges with sum $1/(1-x)$ for $|x| < 1$, and diverges for $|x| \geq 1$, as you saw in Theorem D24 in Unit D3 *Series*. Thus, if $|x| < 1$, then

$$T_n(x) \to f(x) \text{ as } n \to \infty.$$

So in this example we can prove that the polynomials $T_n(x)$ provide better and better approximations to $f(x)$ as n increases, but only if $|x| < 1$. For $|x| \geq 1$, the sequence $(T_n(x))$ does not converge, so increasing the value of n does not in general give a better approximation to $f(x)$.

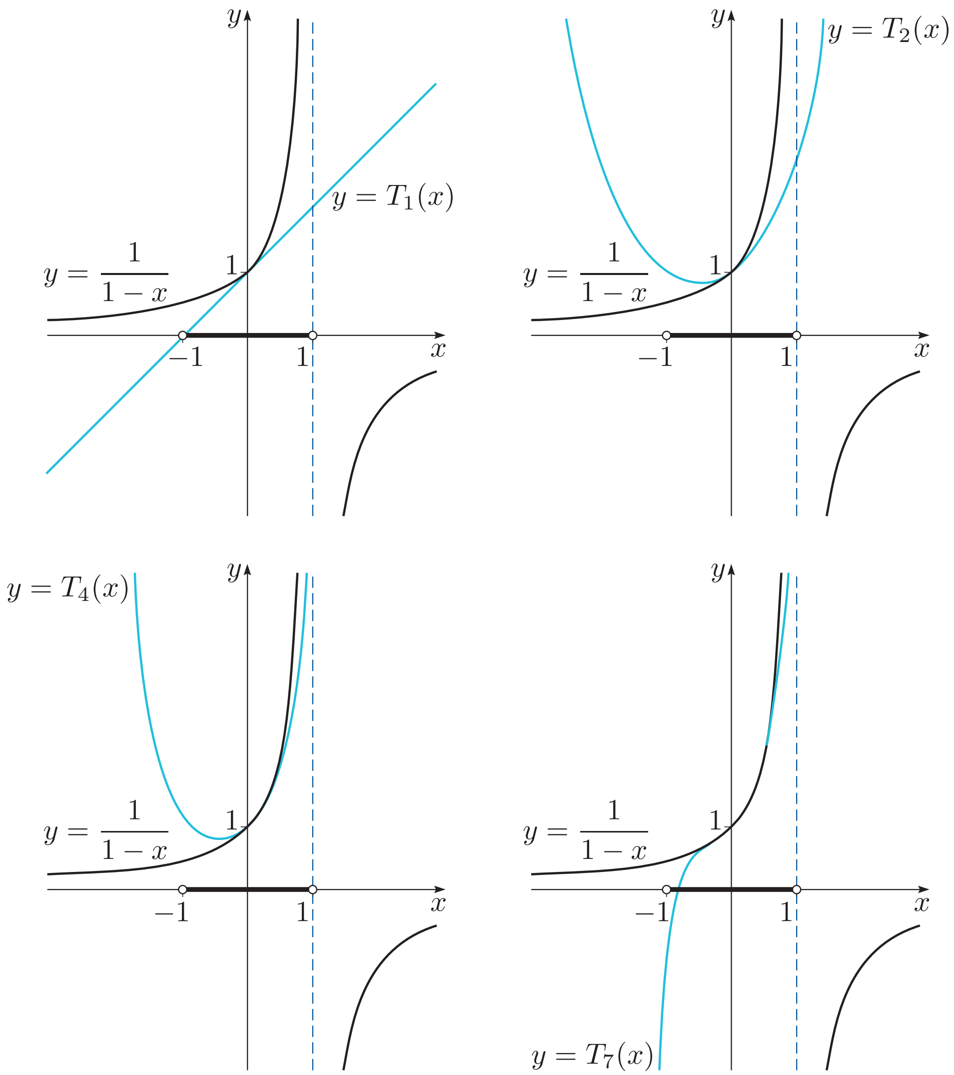


Figure 4 The graphs of Taylor polynomials at 0 for $f(x) = 1/(1-x)$

In later sections we often need general formulas for certain key Taylor polynomials at the point $a = 0$; we ask you to find these in the next exercise.

Exercise F59

Determine the Taylor polynomial of degree n at 0 for each of the following functions.

(a) $f(x) = e^x$ (b) $f(x) = \log(1 + x)$ (c) $f(x) = \cos x$

2 Taylor's Theorem

In this section you will investigate how closely the Taylor polynomials of a function f approximate f and see that, for many functions, f has a representation at the point $x = a$ of the form

$$f(x) = \lim_{n\to\infty} T_n(x) = \sum_{n=0}^{\infty} a_n(x-a)^n.$$

We then say that f is the *sum function* of a *power series*.

2.1 Taylor's Theorem

In Section 1 we showed how to find the Taylor polynomial $T_n(x)$ of degree n at the point a for a function f. This polynomial and its first n derivatives agree with f and its first n derivatives at a, and for larger values of n the polynomial appears to approximate f at points near a. The following fundamental result gives a formula for the error involved in this approximation.

Theorem F63 Taylor's Theorem

Let the function f be $(n+1)$-times differentiable on an open interval containing the points a and x. Then

$$\begin{aligned} f(x) &= f(a) + f'(a)(x-a) + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_n(x) \\ &= T_n(x) + R_n(x), \end{aligned}$$

where $T_n(x)$ is the Taylor polynomial of degree n at a for f and

$$R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1},$$

for some c between a and x.

Remarks

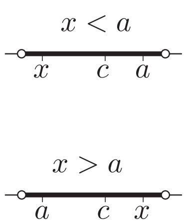


Figure 5 The points a, x and c in Taylor's Theorem

1. Figure 5 shows the possible relative positions of the points a, x and c.
2. Since Taylor's Theorem can be expressed in the form

   $$f(x) = T_n(x) + R_n(x),$$

   we say that $R_n(x)$ is a **remainder term**, or **error term**, involved in approximating $f(x)$ by $T_n(x)$. The formula for $R_n(x)$ involves an 'unknown number' c, so it does not specify the remainder term $R_n(x)$ exactly. Nevertheless, we can often use it to show that $T_n(x)$ is a good approximation to $f(x)$. Note that, as with the notation $T_n(x)$, our notation $R_n(x)$ does not reflect the fact that the remainder also depends on both f and a.

3. When $n = 0$, Taylor's Theorem reduces to

$$f(x) = f(a) + f'(c)(x - a),$$

for some c between a and x; that is,

$$\frac{f(x) - f(a)}{x - a} = f'(c),$$

for some c between a and x. But this is just the Mean Value Theorem that you met in Subsection 4.1 of Unit F2 *Differentiation* (see Figure 6). Thus Taylor's Theorem can be considered as a generalisation of the Mean Value Theorem.

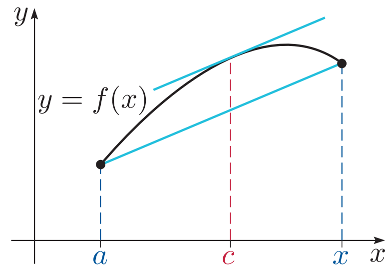


Figure 6 The Mean Value Theorem

4. The form of the remainder $R_n(x)$ given in Taylor's Theorem is actually due to Lagrange. There are other forms, due to Taylor and Cauchy, and also the following neat formula which can be derived by repeated integration by parts (we do not prove this here):

$$R_n(x) = \frac{1}{n!} \int_a^x (x - t)^n f^{(n+1)}(t)\, dt.$$

Note that this form of the remainder does not involve any 'unknown numbers'.

Brook Taylor (1685–1731) was an English mathematician who in 1715 published a slim volume entitled *Methodus incrementorum* (Method of Increments) which included the theorem that now bears his name. Taylor was the first to publish the theorem but he was not the first to discover it. At least five mathematicians anticipated him: James Gregory (1671), Leibniz (1670s), Newton (1691), Johann Bernoulli (1694) and de Moivre (1708). However, Taylor was the first to have appreciated the fundamental significance of the result.

The first explicit expression for the remainder term in Taylor's theorem was provided by Joseph-Louis Lagrange in 1797 in his *Théorie des fonctions analytiques* (Theory of analytic functions), the text in which he attempted to provide a sound foundation for calculus by reducing it to algebra and developing it on the basis of Taylor's Theorem.

Taylor was an accomplished musician and artist, and in his book on linear perspective, which was also first published in 1715, with an improved version in 1719, he enriched the theory of perspective in many respects.

Brook Taylor

In Subsection 4.1 of Unit F2 we proved the Mean Value Theorem using Rolle's Theorem, which you also met in Unit F2. This says that, if f is continuous on the closed interval $[a, b]$ and differentiable on (a, b), and if $f(a) = f(b)$, then there is a point c with $a < c < b$ such that $f'(c) = 0$. We now use Rolle's Theorem to prove Taylor's Theorem. If you are short of time, then you may prefer to skim read this proof.

Proof of Taylor's Theorem In the proof we assume that $x > a$; the proof for $x < a$ is similar.

We consider the function

$$h(t) = f(t) - T_n(t) - A(t-a)^{n+1}, \tag{1}$$

where T_n is the Taylor polynomial of degree n at a for f, and A is a constant chosen so that

$$h(x) = 0. \tag{2}$$

Now, by the definition of T_n,

$$f(a) = T_n(a), \quad f'(a) = T_n'(a), \quad \ldots, \quad f^{(n)}(a) = T_n^{(n)}(a),$$

so

$$h(a) = 0, \quad h'(a) = 0, \quad \ldots, \quad h^{(n)}(a) = 0.$$

Thus the function h is continuous and differentiable on an open interval containing a and x, and $h(a) = 0 = h(x)$. So, by Rolle's Theorem applied to h on the interval $[a, x]$, there is a number c_1 between a and x for which

$$h'(c_1) = 0.$$

Similarly, the function h' is continuous and differentiable on an open interval containing a and c_1, and $h'(a) = 0 = h'(c_1)$. Hence, by Rolle's Theorem applied to h' on the interval $[a, c_1]$, there is a number c_2 between a and c_1 for which

$$h''(c_2) = 0.$$

Applying Rolle's Theorem successively to the functions

$$h'', \ h^{(3)}, \ \ldots, \ h^{(n)},$$

on the intervals

$$[a, c_2], \ [a, c_3], \ \ldots, \ [a, c_n], \quad \text{where } c_2 > c_3 > \cdots > c_n > a,$$

we deduce that there is a number c between a and c_n for which

$$h^{(n+1)}(c) = 0. \tag{3}$$

These points are illustrated in Figure 7.

Figure 7 The points a, c, c_1, c_2, $\ldots$, c_n and x

By repeatedly differentiating equation (1), we obtain

$$h^{(n+1)}(t) = f^{(n+1)}(t) - A(n+1)!. \tag{4}$$

Note that $T_n^{(n+1)}(t) = 0$ since T_n is a polynomial of degree at most n and differentiating such a polynomial n times gives a constant.

From equations (3) and (4), we deduce that

$$f^{(n+1)}(c) - A(n+1)! = 0, \quad \text{so} \quad A = \frac{f^{(n+1)}(c)}{(n+1)!}. \tag{5}$$

Finally, it follows from equations (1), (2) and (5) that

$$f(x) = T_n(x) + A(x-a)^{n+1} = T_n(x) + \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1},$$

as required. ■

Exercise F60

By applying Taylor's Theorem with $n = 3$ to the function $f(x) = \cos x$ at $a = 0$, prove that, for $x \neq 0$,

$$\cos x = 1 - \tfrac{1}{2}x^2 + \frac{\cos c}{4!}\,x^4,$$

where c lies between 0 and x.

(The Taylor polynomial of degree n at 0 for $\cos x$ was found in Exercise F59(c)).

The conclusion of Exercise F60 can be restated as

$$\cos x - (1 - \tfrac{1}{2}x^2) = \frac{\cos c}{4!}x^4,$$

where c lies between 0 and x. Here we do not know the exact value of c, but we do know that $|\cos c| \leq 1$. Thus we can deduce that

$$\left|\cos x - (1 - \tfrac{1}{2}x^2)\right| = \frac{|\cos c|}{4!}|x|^4 \leq \frac{|x|^4}{4!}.$$

In this way, we obtain an explicit **remainder estimate**, or **error bound**, for the approximation of $\cos x$ by $1 - \frac{1}{2}x^2$, which is small when x is near 0.

In general, we can obtain such an estimate for $|f(x) - T_n(x)| = |R_n(x)|$ provided that we have an estimate for $|f^{(n+1)}(c)|$ which is valid for all c between a and x. The following strategy sets out this process.

Strategy F11

To show that the Taylor polynomial T_n at a for f approximates f to a certain accuracy at a point $x \neq a$, do the following.

1. Obtain a formula for $f^{(n+1)}$.
2. Determine a number M such that
$$|f^{(n+1)}(c)| \leq M, \quad \text{for all } c \text{ between } a \text{ and } x.$$
3. Write down and simplify the remainder estimate
$$|f(x) - T_n(x)| = |R_n(x)| \leq \frac{M}{(n+1)!}|x-a|^{n+1}.$$

Note that in step 2 we can use any convenient value for M, preferably not too large. Sometimes we can determine the maximum value of $|f^{(n+1)}(c)|$ for c in the closed interval with endpoints a and x, and often this maximum value is taken when c is equal to either a or x. Usually, however, any 'good enough' upper bound for $|f^{(n+1)}(c)|$ will do.

Worked Exercise F43

(a) Write down the Taylor polynomial $T_3(x)$ at $a = 0$ for $f(x) = \sin x$.

(b) Use Taylor's Theorem to show that $|\sin(0.1) - T_3(0.1)| < 5 \times 10^{-6}$.

(c) Hence calculate $\sin(0.1)$ to four decimal places.

Solution

(a) For $f(x) = \sin x$ and $a = 0$, we have

$$T_3(x) = x - \tfrac{1}{6}x^3.$$

This expression for $T_3(x)$ was derived in Worked Exercise F41(a).

(b) We use Strategy F11 with $a = 0$, $x = 0.1$ and $n = 3$.

1. First, $f^{(4)}(x) = \sin x$.
2. Thus

$$|f^{(4)}(c)| = |\sin c| \leq 1, \quad \text{for } c \in [0, 0.1],$$

so we can take $M = 1$.

With care a smaller value for M can be obtained. For example, using the Sine Inequality (Theorem D45 in Unit D4 *Continuity*), we have

$$|\sin c| \leq |c| \leq 0.1, \quad \text{for } c \in [0, 0.1].$$

However, here we take $M = 1$.

3. Using the remainder estimate $\dfrac{M}{(n+1)!}|x-a|^{n+1}$, we therefore obtain

$$\begin{aligned} |\sin(0.1) - T_3(0.1)| &= |R_3(0.1)| \\ &\leq \frac{M}{(3+1)!}\,|x-a|^{3+1} \\ &= \frac{1}{4!}\,|0.1-0|^4 \\ &= 0.000\,004\,1\overline{6} \\ &< 5 \times 10^{-6}, \end{aligned}$$

as required.

(c) By part (a),

$$\begin{aligned} T_3(0.1) &= 0.1 - \tfrac{1}{6} \times 10^{-3} \\ &= 0.1 - 0.000\,166\,66\ldots = 0.099\,833\,33\ldots. \end{aligned}$$

By part (b),

$$|\sin(0.1) - T_3(0.1)| = |R_3(0.1)| < 5 \times 10^{-6}.$$

Hence

$$0.099\,828\,33\ldots < \sin(0.1) < 0.099\,838\,33\ldots,$$

so

$$\sin(0.1) = 0.0998 \quad \text{(to 4 d.p.)}.$$

Exercise F61

(a) Write down the Taylor polynomial $T_2(x)$ at $a = 0$ for $f(x) = \log(1 + x)$, using the solution to Exercise F59(b).

(b) Use Taylor's Theorem to show that $|\log(1.02) - T_2(0.02)| < 3 \times 10^{-6}$.

(c) Hence calculate $\log(1.02)$ to four decimal places.

Our next strategy shows how to use Taylor's Theorem to obtain an approximation to f which holds at all points of an interval.

Strategy F12

To show that the Taylor polynomial T_n at a for f approximates f to a certain accuracy throughout an interval I of the form $[a, a + r]$, $[a - r, a]$ or $[a - r, a + r]$, where $r > 0$, do the following.

1. Obtain a formula for $f^{(n+1)}$.
2. Determine a number M such that
$$|f^{(n+1)}(c)| \leq M, \quad \text{for all } c \in I.$$
3. Write down and simplify the remainder estimate
$$|f(x) - T_n(x)| = |R_n(x)| \leq \frac{M}{(n+1)!} r^{n+1}, \quad \text{for all } x \in I.$$

Strategy F12 is obtained from Strategy F11 by replacing $|x - a|$ by r, since $|x - a| \leq r$ for all values of x in the interval I.

By applying Strategy F12 to the situation in Worked Exercise F43, we can show that the remainder estimate at the point $x = 0.1$ in fact holds over the whole interval $[0, 0.1]$. Here we have $r = 0.1$, $M = 1$ and $n = 3$, so

$$|\sin x - T_3(x)| \leq \frac{1}{4!} 0.1^4 < 5 \times 10^{-6}, \quad \text{for all } x \in [0, 0.1].$$

Here is another worked exercise.

Worked Exercise F44

(a) Calculate the Taylor polynomial $T_3(x)$ at 1 for $f(x) = 1/(x+2)$.

(b) Show that $T_3(x)$ approximates $f(x)$ with an error less than 5×10^{-3} on the interval $[1, 2]$.

Solution

(a) For this function,

$$\begin{aligned} f(x) &= 1/(x+2), & f(1) &= 1/3 \\ f'(x) &= -1/(x+2)^2, & f'(1) &= -1/9 \\ f''(x) &= 2/(x+2)^3, & f''(1) &= 2/27 \\ f^{(3)}(x) &= -6/(x+2)^4, & f^{(3)}(1) &= -2/27. \end{aligned}$$

Hence the Taylor polynomial of degree 3 at 1 for f is

$$T_3(x) = \tfrac{1}{3} - \tfrac{1}{9}(x-1) + \tfrac{1}{27}(x-1)^2 - \tfrac{1}{81}(x-1)^3.$$

(b) We use Strategy F12 with $I = [1, 2]$, $a = 1$, $r = 1$ and $n = 3$.

1. First, $f^{(4)}(x) = \dfrac{24}{(x+2)^5}$.

2. Thus
$$|f^{(4)}(c)| = \frac{24}{(c+2)^5} \leq \frac{24}{3^5}, \quad \text{for } c \in [1, 2],$$
so we can take $M = 24/3^5$.
Here we have used the fact that, since $c \geq 1$, we have $c + 2 \geq 1 + 2 = 3$.

3. Using the remainder estimate $\dfrac{M}{(n+1)!}\, r^{n+1}$, we obtain
$$\begin{aligned} |f(x) - T_3(x)| &= |R_3(x)| \\ &\leq \frac{M}{(3+1)!}\, r^{3+1} \\ &= \frac{1}{4!} \times \frac{24}{3^5} \times 1^4 \\ &= \frac{1}{3^5} = 0.0041\ldots, \quad \text{for } x \in [1, 2]. \end{aligned}$$

Thus $T_3(x)$ approximates $f(x)$ with an error less than 5×10^{-3} on $[1, 2]$.

Exercise F62

(a) Calculate the Taylor polynomial $T_4(x)$ at π for the function $f(x) = \cos x$.

(b) Show that $T_4(x)$ approximates $f(x)$ with an error less than 3×10^{-3} on the interval $[3\pi/4, 5\pi/4]$.

2.2 Taylor series

From Taylor's Theorem we know that if a function f can be differentiated as often as we want on an open interval containing the points a and x, then

$$f(x) = T_n(x) + R_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + R_n(x),$$

for $n = 0, 1, 2, \ldots$, where

$$R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1},$$

for some c between a and x. So if the error terms $R_n(x)$ tend to zero, then the Taylor polynomials $T_n(x)$ tend to $f(x)$ as n tends to infinity, and we can write f as an infinite series. Thus we have the following result.

Theorem F64

Let f have derivatives of all orders on an open interval containing the points a and x. If

$$R_n(x) \to 0 \text{ as } n \to \infty,$$

then

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

Note that, for $x = a$ and $n = 0$, the series in the statement of Theorem F64 involves the expression 0^0. By convention, we take $0^0 = 1$ in such series.

It follows from Theorem F64 that, if $R_n(x) \to 0$ as $n \to \infty$, then we can express $f(x)$ as a series whose terms involve powers of $(x-a)$.

Definition

Let f have derivatives of all orders at the point a. The **Taylor series at a for f** is

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots.$$

If x is a point for which the Taylor series for f has the sum $f(x)$ given in the statement of Theorem F64 (that is, when $R_n(x) \to 0$ as $n \to \infty$), then we say that the Taylor series is **valid** at the point x. Any set of values of x for which a Taylor series is valid is called a **range of validity** for the Taylor series. On any such range of validity, the function f is the **sum function** of the Taylor series.

We can use Theorem F64 to obtain the following basic Taylor series. In each case we have indicated the largest possible range of validity.

Theorem F65 Basic Taylor series at 0

(a) $\dfrac{1}{1-x} = 1 + x + x^2 + \cdots = \displaystyle\sum_{n=0}^{\infty} x^n$, for $|x| < 1$.

(b) $\sin x = x - \dfrac{x^3}{3!} + \dfrac{x^5}{5!} - \cdots = \displaystyle\sum_{n=0}^{\infty} \dfrac{(-1)^n x^{2n+1}}{(2n+1)!}$, for $x \in \mathbb{R}$.

(c) $\cos x = 1 - \dfrac{x^2}{2!} + \dfrac{x^4}{4!} - \cdots = \displaystyle\sum_{n=0}^{\infty} \dfrac{(-1)^n x^{2n}}{(2n)!}$, for $x \in \mathbb{R}$.

(d) $e^x = 1 + x + \dfrac{x^2}{2!} + \cdots = \displaystyle\sum_{n=0}^{\infty} \dfrac{x^n}{n!}$, for $x \in \mathbb{R}$.

(e) $\log(1+x) = x - \dfrac{x^2}{2} + \dfrac{x^3}{3} - \cdots = \displaystyle\sum_{n=1}^{\infty} \dfrac{(-1)^{n+1} x^n}{n}$, for $-1 < x \leq 1$.

Remarks

1. Taking $x = 1$ in Theorem F65(e), we obtain the unexpected sum

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \tfrac{1}{2} + \tfrac{1}{3} - \tfrac{1}{4} + \cdots = \log 2.$$

This series is known as the **alternating harmonic series**.

2. In Subsection 4.1 of Unit D3 we defined the exponential function as

$$e^x = \begin{cases} \displaystyle\sum_{n=0}^{\infty} \frac{x^n}{n!}, & x \geq 0, \\ (e^{-x})^{-1}, & x < 0. \end{cases}$$

Theorem F65(d) shows that e^x is the sum function of the series $\displaystyle\sum_{n=0}^{\infty} x^n/n!$ for *all* x, not just for $x \geq 0$. Note that some texts define $e^x = \exp(x)$ using this power series.

3. In this module, $\sin x$ and $\cos x$ are defined in terms of a right-angled triangle. Theorem F65 shows that $\sin x$ and $\cos x$ can be represented by power series for all $x \in \mathbb{R}$, and some texts use these series to define $\sin x$ and $\cos x$ in a way that does not depend on geometric ideas.

Proof of Theorem F65

(a) Let $f(x) = 1/(1-x)$. You saw in Subsection 1.2 that the Taylor polynomial of degree n at 0 for f is

$$T_n(x) = 1 + x + \cdots + x^n = \sum_{k=0}^{n} x^k.$$

Now $1 + x + x^2 + \cdots$ is a geometric series with initial term 1 and common ratio x, which has sum $1/(1-x)$ for $|x| < 1$, as you saw in Theorem D24 in Unit D3. Thus the result follows.

(b) Let $f(x) = \sin x$. You saw in Worked Exercise F42 that the Taylor polynomial of degree n at 0 for f is

$$T_n(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + \frac{(-1)^m x^{2m+1}}{(2m+1)!} = \sum_{k=0}^{m} \frac{(-1)^k x^{2k+1}}{(2k+1)!},$$

where $n = 2m+1$ or $n = 2m+2$.

By Taylor's Theorem, we have

$$\sin x = T_n(x) + R_n(x), \quad \text{where } R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}\, x^{n+1},$$

for some c between 0 and x. Since

$$f^{(n+1)}(x) = \pm \sin x \text{ or } \pm \cos x,$$

we have $|f^{(n+1)}(c)| \leq 1$, so

$$|R_n(x)| \leq \frac{|x|^{n+1}}{(n+1)!} \to 0 \text{ as } n \to \infty.$$

Here we have used the Squeeze Rule for null sequences and the fact that $(x^n/n!)$ is a basic null sequence; see Subsection 2.3 of Unit D2 *Sequences*.

Hence the result follows.

(c) The proof of part (c) is similar to that of part (b), so we omit the details.

(d) Let $f(x) = e^x$. You saw in the solution to Exercise F59(a) that the Taylor polynomial of degree n at 0 for f is

$$T_n(x) = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} = \sum_{k=0}^{n} \frac{x^k}{k!}.$$

By Taylor's Theorem, we have

$$e^x = T_n(x) + R_n(x),$$

where

$$R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}\, x^{n+1} = e^c \frac{x^{n+1}}{(n+1)!},$$

for some c between 0 and x. Now $e^c \leq e^{|x|}$ and $(x^{n+1}/(n+1)!)$ is a null sequence.

The value of c depends on n, but it always lies between 0 and x.

Hence $R_n(x) \to 0$ as $n \to \infty$, by the Squeeze Rule for null sequences, so the result follows.

(e) Let $f(x) = \log(1 + x)$. You saw in the solution to Exercise F59(b) that the Taylor polynomial of degree n at 0 for f for $n \geq 1$ is

$$T_n(x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots + \frac{(-1)^{n+1}x^n}{n} = \sum_{k=1}^{n} \frac{(-1)^{k+1}x^k}{k}.$$

By Taylor's Theorem, we have

$$\log(1 + x) = T_n(x) + R_n(x), \quad \text{where } R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}\, x^{n+1},$$

for some c between 0 and x.

Now suppose that $0 < x \leq 1$. Since, by the solution to Exercise F59(b),

$$f^{(n+1)}(x) = \frac{(-1)^{n+2}n!}{(1+x)^{n+1}},$$

we have

$$\begin{aligned} |R_n(x)| &= \frac{1}{(n+1)!} \times \frac{n!}{(1+c)^{n+1}}\, |x|^{n+1} \\ &= \frac{|x|^{n+1}}{(n+1)(1+c)^{n+1}} \leq \frac{1}{n} \to 0 \text{ as } n \to \infty. \end{aligned}$$

Hence

$$\log(1 + x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}x^n}{n}, \quad \text{for } 0 \leq x \leq 1.$$

Here we have used the fact that $0 < c < x \leq 1$, so

$$|x|^{n+1} \leq 1 \quad \text{and} \quad 1 + c > 1.$$

We can include $x = 0$ in the range of validity because both $\log(1 + x)$ and the sum function are zero when $x = 0$.

The proof that this Taylor series is also valid for $-1 < x < 0$ does not follow from the above form of $R_n(x)$; we ask you to prove it later, in Exercise F72.

3 Convergence of power series

So far in this unit you have seen how a function f can often be approximated near a point a by means of a Taylor series, which is an infinite sum of powers of $(x - a)$. In this section you will study the behaviour of such *power series* in their own right, and consider functions which are *defined* by power series.

3.1 Radius of convergence

We begin with a formal definition of a power series.

Definitions

Let $a \in \mathbb{R}$, $x \in \mathbb{R}$ and $a_n \in \mathbb{R}$, $n = 0, 1, 2, \ldots$. Then the expression

$$\sum_{n=0}^{\infty} a_n(x-a)^n = a_0 + a_1(x-a) + a_2(x-a)^2 + \cdots$$

is called a **power series at a in x**, with **coefficients** a_n. We call a the **centre** of the power series.

Notice that in the definition of a power series, we think of a as a constant and x as a variable.

In Section 2 you saw that certain standard functions can be expressed as the sum functions of their Taylor series; for example,

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n, \quad \text{for } |x| < 1,$$

and

$$\log(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}x^n}{n}, \quad \text{for } -1 < x \le 1,$$

and both these series are power series at 0 in x. All Taylor series are examples of power series.

On the other hand, we can consider power series in their own right and use these to *define* functions. For example, you saw in Subsection 2.2 that we could have defined the exponential function by the formula

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad (x \in \mathbb{R}).$$

Another example is the *Bessel function*

$$J_0(x) = \sum_{n=0}^{\infty} \frac{(-1)^n (x/2)^{2n}}{(n!)^2} \quad (x \in \mathbb{R}),$$

which arises in connection with the vibration of a circular drum.

In each of the above examples (where $a = 0$), the power series converges on an interval with centre a. The next result shows that this property is true for all power series; it is illustrated in Figure 8. We give the proof of this result in Subsection 3.2.

Figure 8 The radius of convergence

Theorem F66 Radius of Convergence Theorem

For a given power series $\sum_{n=0}^{\infty} a_n(x-a)^n$, exactly one of the following possibilities occurs.

(a) The series converges only for $x = a$.

(b) The series converges for all x.

(c) There is a number $R > 0$ such that

$$\sum_{n=0}^{\infty} a_n(x-a)^n \text{ converges if } |x-a| < R$$

and

$$\sum_{n=0}^{\infty} a_n(x-a)^n \text{ diverges if } |x-a| > R.$$

Moreover, in parts (a), (b) and (c) the series converges absolutely on the specified sets of convergence.

Although it is usually sufficient to know that a series is convergent, the stronger result about absolute convergence is sometimes useful (for example, we will use it when we prove the Differentiation Rule in Subsection 4.2). Remember that a series $\sum a_n$ converges absolutely if the series $\sum |a_n|$ converges. We showed in Theorem D34 in Unit D3 that every absolutely convergent series is convergent.

You have already met the following examples of the three possibilities in Theorem F66 (see Theorem F65(d) for (b) and Unit D3 for (a) and (c)):

(a) $\sum_{n=0}^{\infty} n!\, x^n$ converges only for $x = 0$

(b) $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ converges for all x

(c) $\sum_{n=0}^{\infty} x^n$ converges if $|x| < 1$ and diverges if $|x| > 1$, so $R = 1$.

The positive number R in Theorem F66(c) is called the **radius of convergence** of the power series because the power series converges at those points whose distance from the centre a is less than R, and diverges at those points whose distance from a is greater than R. (Power series can also be defined for complex variables, in which case the points whose distance from the centre is less than R form a disc of radius R.) We extend the definition of the radius of convergence to the cases of Theorem F66(a) and (b) by writing

$R = 0$ if the power series converges only for $x = a$

and

$R = \infty$ if the power series converges for all x.

In this last case R is used as a symbol, not a real number.

Theorem F66(c) makes no assertion about the behaviour of the power series at the endpoints of the interval $(a - R, a + R)$; in fact, a power series may converge at both endpoints, neither endpoint or exactly one endpoint, as you will see in Worked Exercise F46.

The **interval of convergence** of the power series is the interval $(a - R, a + R)$, together with any endpoints of this interval at which the power series converges.

Figure 9 illustrates the various possible types of interval of convergence of $\sum_{n=0}^{\infty} a_n(x-a)^n$.

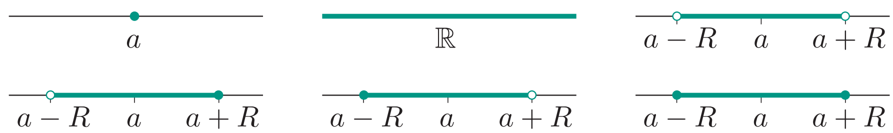


Figure 9 Possible types of intervals of convergence

Theorem F66 tells us that each power series has a radius of convergence R, but it does not tell us how to find R. However, a power series is a particular type of series, so the convergence tests for series from Unit D3 can be applied.

We can find the radius of convergence of many power series by using the following version of the Ratio Test for series from Subsection 2.1 of Unit D3.

Theorem F67 Ratio Test for power series

Suppose that $\sum_{n=0}^{\infty} a_n(x-a)^n$ is a power series with radius of convergence R, and that

$$\left|\frac{a_{n+1}}{a_n}\right| \to L \text{ as } n \to \infty.$$

(a) If L is ∞, then $R = 0$.

(b) If $L = 0$, then $R = \infty$.

(c) If $L > 0$, then $R = 1/L$.

Proof We give this proof only in the case that $a = 0$. The proof of the general case is similar. It follows from the statement about absolute convergence in Theorem F66 that it is sufficient to consider the convergence of the series $\sum_{n=0}^{\infty} |a_n x^n|$ as this is convergent precisely when $\sum_{n=0}^{\infty} a_n x^n$ is convergent. This enables us to base the proof on the Ratio Test for series which can only be applied to a series of positive terms.

(a) Suppose that

$$\left|\frac{a_{n+1}}{a_n}\right| \to \infty \text{ as } n \to \infty.$$

Then, for $x \neq 0$,

$$\frac{|a_{n+1}x^{n+1}|}{|a_n x^n|} = \left|\frac{a_{n+1}}{a_n}\right| |x| \to \infty \text{ as } n \to \infty,$$

so $\sum_{n=0}^{\infty} |a_n x^n|$ is divergent, by the Ratio Test for series.

Thus the series $\sum_{n=0}^{\infty} a_n x^n$ converges only for $x = 0$, so $R = 0$.

(b) Now suppose that

$$\left|\frac{a_{n+1}}{a_n}\right| \to 0 \text{ as } n \to \infty.$$

Then, for $x \neq 0$,

$$\frac{|a_{n+1}x^{n+1}|}{|a_n x^n|} = \left|\frac{a_{n+1}}{a_n}\right| |x| \to 0 \times |x| = 0 < 1 \text{ as } n \to \infty,$$

so $\sum_{n=0}^{\infty} |a_n x^n|$ is convergent, by the Ratio Test for series.

Thus the series $\sum_{n=0}^{\infty} a_n x^n$ converges for all $x \in \mathbb{R}$, so $R = \infty$.

(c) Finally, suppose that

$$\left|\frac{a_{n+1}}{a_n}\right| \to L \text{ as } n \to \infty,$$

where $L > 0$.

If $|x| > 1/L$, then

$$\frac{|a_{n+1}x^{n+1}|}{|a_n x^n|} = \left|\frac{a_{n+1}}{a_n}\right| |x| \to L|x| > 1 \text{ as } n \to \infty,$$

so $\sum_{n=0}^{\infty} |a_n x^n|$ is divergent, by the Ratio Test for series. Thus if $|x| > L$, then the series $\sum_{n=0}^{\infty} a_n x^n$ is not absolutely convergent and hence is not convergent, so it follows that $R \leq 1/L$.

However, if $0 < |x| < 1/L$, then

$$\frac{|a_{n+1}x^{n+1}|}{|a_n x^n|} = \left|\frac{a_{n+1}}{a_n}\right| |x| \to L|x| < 1 \text{ as } n \to \infty,$$

so $\sum_{n=0}^{\infty} |a_n x^n|$ is convergent, by the Ratio Test for series. Thus if $|x| < L$, then the series $\sum_{n=0}^{\infty} a_n x^n$ is absolutely convergent and hence, by Theorem F66, is convergent, from which it follows that $R \geq 1/L$.

Taken together, these results show that $R = 1/L$, which completes the proof. ■

Worked Exercise F45

Determine the radius of convergence of each of the following power series.

(a) $\displaystyle\sum_{n=0}^{\infty} \frac{(x+2)^n}{n!}$ (b) $\displaystyle\sum_{n=0}^{\infty} \frac{n^n(x-1)^n}{n!}$

Solution

(a) Here $a_n = 1/n!$, for $n = 0, 1, 2, \ldots$, so

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{1}{(n+1)!} \times \frac{n!}{1} = \frac{1}{n+1} \to 0 \text{ as } n \to \infty.$$

Hence, by the Ratio Test, $R = \infty$.

Thus this power series converges for all x.

(b) Here $a_n = n^n/n!$, for $n = 0, 1, 2, \ldots$, so

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{(n+1)^{n+1}}{(n+1)!} \times \frac{n!}{n^n}$$
$$= \frac{(n+1)^n}{n^n} = \left(1 + \frac{1}{n}\right)^n \to e \text{ as } n \to \infty.$$

You met this limit in Subsection 5.3 of Unit D2.

Hence, by the Ratio Test, the radius of convergence is $R = 1/e$.

Thus this power series converges for $|x - 1| < 1/e$, and diverges for $|x - 1| > 1/e$.

Exercise F63

Determine the radius of convergence of each of the following power series.

(a) $\displaystyle\sum_{n=0}^{\infty} (2^n + 4^n)x^n$ (b) $\displaystyle\sum_{n=1}^{\infty} \frac{(n!)^2}{(2n)!}x^n$ (c) $\displaystyle\sum_{n=0}^{\infty} (n + 2^{-n})(x-1)^n$

(d) $\displaystyle\sum_{n=1}^{\infty} n^n x^n$

The next exercise concerns a power series which plays an important role in Section 4, where we will use it to prove a generalised binomial theorem. In this power series α can be any real number, but if $\alpha \in \{0, 1, 2, \ldots\}$ then the series has only finitely many non-zero terms and thus converges for all values of x. The exercise asks you to determine the radius of convergence of the power series for other values of α.

Exercise F64

Determine the radius of convergence of the power series

$$1 + \alpha x + \frac{\alpha(\alpha-1)}{2!}x^2 + \cdots + \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}x^n + \cdots,$$

where $\alpha \neq 0, 1, 2, \ldots$.

The Ratio Test gives an open interval on which a power series converges. To determine the full interval of convergence of a power series with finite non-zero radius of convergence, we need to use other tests to find the behaviour at the interval endpoints.

Strategy F13

To find the interval of convergence of the power series $\displaystyle\sum_{n=0}^{\infty} a_n(x-a)^n$, do the following.

1. Use the Ratio Test for power series to find the radius of convergence R.
2. If R is finite and non-zero, use other tests for series to determine the behaviour of the power series at the endpoints of the interval $(a-R, a+R)$.

You met a general strategy for applying the various tests for convergence or divergence of a series in Subsection 3.3 of Unit D3 (Strategy D13, which you can also find in the module Handbook). You may find it helpful to refer to this general strategy when applying step 2 of Strategy F13.

Note in particular that we can use Strategy F13 to establish the largest possible range of validity of a Taylor series. Since a Taylor series is a power series, its largest possible range of validity is the interval of convergence obtained by using Strategy F13.

In the following worked exercise and exercise, each of the power series has coefficient $a_0 = 0$, so the sum starts at $n = 1$.

Worked Exercise F46

Determine the interval of convergence of each of the following power series.

(a) $\displaystyle\sum_{n=1}^{\infty} x^n$ (b) $\displaystyle\sum_{n=1}^{\infty} \frac{x^n}{n}$ (c) $\displaystyle\sum_{n=1}^{\infty} \frac{(x-3)^n}{2^n n^2}$

Solution

In each case, we apply Strategy F13.

(a) Here $a_n = 1$, for $n = 1, 2, \ldots$.

1. Since
$$\left|\frac{a_{n+1}}{a_n}\right| = 1, \quad \text{for } n = 1, 2, \ldots,$$
we have $R = 1$, by the Ratio Test. Thus (by the Radius of Convergence Theorem) this power series
   - converges for $-1 < x < 1$,
   - diverges for $x > 1$ and $x < -1$.

2. If $x = 1$, then the power series is $\displaystyle\sum_{n=1}^{\infty} 1^n$, which is divergent by the Non-null Test.

   If $x = -1$, then the power series is $\displaystyle\sum_{n=1}^{\infty} (-1)^n$, which is also divergent by the Non-null Test.

Hence the interval of convergence is $(-1, 1)$.

(b) Here $a_n = 1/n$, for $n = 1, 2, \ldots$.

1. Since
$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{1}{n+1} \times \frac{n}{1} = \frac{1}{1+1/n} \to 1 \text{ as } n \to \infty,$$
we have $R = 1$, by the Ratio Test. Thus this power series
   - converges for $-1 < x < 1$,
   - diverges for $x > 1$ and $x < -1$.

2. If $x = 1$, then the power series is $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n}$, which is a basic divergent series.

   If $x = -1$, then the power series is $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$, which is convergent by the Alternating Test.

Hence the interval of convergence is $[-1, 1)$.

(c) Here $a_n = 1/(2^n n^2)$, for $n = 1, 2, \ldots$.

1. Since
$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{1}{2^{n+1}(n+1)^2} \times \frac{2^n n^2}{1}$$
$$= \frac{1}{2(1+1/n)^2} \to \frac{1}{2} \text{ as } n \to \infty,$$
we have $R = 2$, by the Ratio Test. Since $a = 3$, this power series
   - converges for $1 < x < 5$,
   - diverges for $x > 5$ and $x < 1$.

   This follows because
$$|x-3| < 2 \iff -2 < x-3 < 2$$
$$\iff 1 < x < 5,$$
by the rules for rearranging inequalities.

2. If $x = 5$, then the power series is
$$\sum_{n=1}^{\infty} \frac{1}{2^n n^2}(5-3)^n = \sum_{n=1}^{\infty} \frac{1}{n^2},$$
which is a basic convergent series.
If $x = 1$, then the power series is
$$\sum_{n=1}^{\infty} \frac{1}{2^n n^2}(1-3)^n = \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2},$$
which is convergent by the Absolute Convergence Test.

Hence the interval of convergence is $[1, 5]$.

Exercise F65

Determine the interval of convergence of each of the following power series.

(a) $\displaystyle\sum_{n=1}^{\infty} nx^n$ (b) $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n}{n3^n}(x-5)^n$

3.2 Proof of the Radius of Convergence Theorem (optional)

In what follows, we often use without reference the fact that an absolutely convergent series is convergent.

To prove the Radius of Convergence Theorem, we need the following preliminary result.

Lemma F68

If the power series $\sum_{n=0}^{\infty} a_n x^n$ converges for some $x_0 \neq 0$, then it is absolutely convergent on the interval $(-|x_0|, |x_0|)$.

Figure 10 illustrates the statement of the lemma in the two possible cases $x_0 > 0$ and $x_0 < 0$.

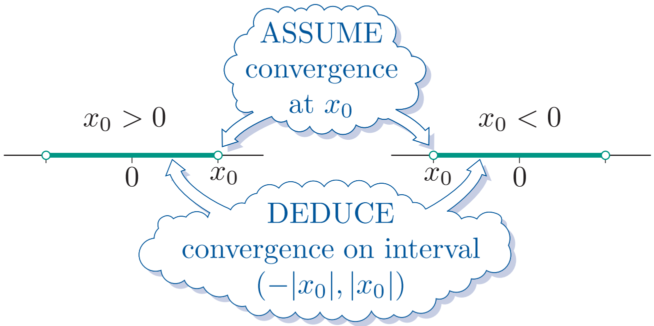


Figure 10 The two cases of Lemma F68

Proof of Lemma F68 First we write $r = |x_0|$. Since the series $\sum_{n=0}^{\infty} a_n x_0^n$ is convergent, the sequence $(a_n x_0^n)$ is null by Theorem D27 in Unit D3, and hence there is a number K such that

$$|a_n| r^n = |a_n x_0^n| \leq K, \quad \text{for } n = 0, 1, 2, \ldots. \tag{6}$$

Suppose that $|x| < r$. To prove that $\sum_{n=0}^{\infty} a_n x^n$ is absolutely convergent, we write

$$a_n x^n = a_n r^n \left(\frac{x}{r}\right)^n.$$

Then, by inequality (6),

$$|a_n x^n| = |a_n| r^n \left|\frac{x}{r}\right|^n \leq K \left(\frac{|x|}{r}\right)^n.$$

Since $|x| < r$, we have $|x|/r < 1$, so the geometric series $\sum_{n=0}^{\infty} K \left(\frac{|x|}{r}\right)^n$ is convergent. Hence, by the Comparison Test (Theorem D30 in Unit D3), $\sum_{n=0}^{\infty} |a_n x^n|$ is convergent, as required. ■

We can now give the proof of the Radius of Convergence Theorem, which we first state again for convenience.

Theorem F66 Radius of Convergence Theorem

For a given power series $\displaystyle\sum_{n=0}^{\infty} a_n(x-a)^n$, exactly one of the following possibilities occurs.

(a) The series converges only for $x = a$.

(b) The series converges for all x.

(c) There is a number $R > 0$ such that

$$\sum_{n=0}^{\infty} a_n(x-a)^n \text{ converges if } |x-a| < R$$

and

$$\sum_{n=0}^{\infty} a_n(x-a)^n \text{ diverges if } |x-a| > R.$$

Moreover, in parts (a), (b) and (c) the series converges absolutely on the specified sets of convergence.

Proof We give the proof only in the case $a = 0$. The proof of the general case is similar.

First we define the set

$$E = \left\{ x \in \mathbb{R} : \sum_{n=0}^{\infty} a_n x^n \text{ is convergent} \right\}.$$

If $E = \{0\}$, then possibility (a) holds.

If E is unbounded, then for every $x \in \mathbb{R}$ there exists some $x_0 \in E$ such that $|x| < |x_0|$. Thus the series $\displaystyle\sum_{n=0}^{\infty} a_n x^n$ is absolutely convergent, by Lemma F68. Since this is true for *every* $x \in \mathbb{R}$, it follows that possibility (b) holds.

Otherwise, the set E is bounded and contains a point $x_0 \neq 0$. Then $(-|x_0|, |x_0|) \subseteq E$, by Lemma F68, so $\sup E \geq |x_0|$. We define $R = \sup E$, where R is the radius of convergence.

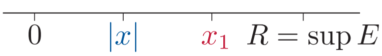


Figure 11 The case when $|x| < R$

If $|x| < R$, then we can find $x_1 \in E$ such that $|x| < x_1$; see Figure 11. Thus, by Lemma F68, the series $\displaystyle\sum_{n=0}^{\infty} a_n x^n$ is absolutely convergent.

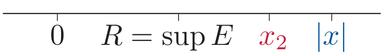


Figure 12 The case when $|x| > R$

If $|x| > R$, then we can find $x_2 > R$ such that $|x| > x_2$; see Figure 12. So $\displaystyle\sum_{n=0}^{\infty} a_n x^n$ is divergent (since if $\displaystyle\sum_{n=0}^{\infty} a_n x^n$ is convergent, then $\displaystyle\sum_{n=0}^{\infty} a_n x_2^n$ is convergent, by Lemma F68). Hence possibility (c) holds.

The above arguments show that in each case the given power series is not just convergent but *absolutely* convergent at each interior point of its interval of convergence.

This completes the proof. ■

4 Manipulating Taylor series

In Section 2 you saw that many functions can be represented by a Taylor series; if there exists $R > 0$ such that

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n,$$

for $|x - a| < R$, then the expression on the right is the Taylor series at a for f.

In this section you will meet several rules which enable us to obtain 'new Taylor series from old', and so build on the list of basic Taylor series given in Theorem F65. Some of the rules for manipulating Taylor series are similar to the corresponding rules for continuous or differentiable functions.

We begin by noting that we can obtain new Taylor series from old by replacing x with another expression. For example, you have already seen the following Taylor series at 0:

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots = \sum_{n=0}^{\infty} x^n, \quad \text{for } |x| < 1. \tag{7}$$

This power series has radius of convergence 1. Since $|-x| < 1$ if and only if $|x| < 1$, we can deduce from equation (7) that

$$\frac{1}{1+x} = 1 - x + x^2 - \cdots = \sum_{n=0}^{\infty} (-1)^n x^n, \quad \text{for } |x| < 1,$$

and this power series also has radius of convergence 1.

Similarly, since $|3x^2| < 1$ if and only if $|x| < 1/\sqrt{3}$, we can deduce from equation (7) that

$$\begin{aligned}\frac{1}{1-3x^2} &= 1 + 3x^2 + (3x^2)^2 + (3x^2)^3 + \cdots \\ &= \sum_{n=0}^{\infty} 3^n x^{2n}, \quad \text{for } |x| < 1/\sqrt{3},\end{aligned}$$

and this power series has radius of convergence $1/\sqrt{3}$.

4.1 The Combination Rules and the Power Rule

We now study the Combination Rules for Taylor series.

Theorem F69 Combination Rules for Taylor series

Let f and g be functions that can both be represented by a Taylor series at a, and suppose that

$$f(x) = \sum_{n=0}^{\infty} a_n (x-a)^n, \text{ for } |x-a| < R,$$

$$g(x) = \sum_{n=0}^{\infty} b_n (x-a)^n, \text{ for } |x-a| < R'.$$

Then the following hold for $r = \min\{R, R'\}$ and $\lambda \in \mathbb{R}$:

Sum Rule $$(f+g)(x) = \sum_{n=0}^{\infty} (a_n + b_n)(x-a)^n, \quad \text{for } |x-a| < r$$

Multiple Rule $$\lambda f(x) = \sum_{n=0}^{\infty} \lambda a_n (x-a)^n, \text{ for } |x-a| < R.$$

Remarks

1. The Sum and Multiple Rules for Taylor series are simply special cases of the Sum and Multiple Rules for general convergent series (see Theorem D25 in Unit D3.)
2. The radius of convergence of the Taylor series for $f+g$ may be larger than $r = \min\{R, R'\}$: Theorem F69 simply asserts that it must be *at least* r. For example, we can use the standard geometric series and the Sum and Multiple Rules to verify that the Taylor series at 0 for the functions $f(x) = \dfrac{1}{1-x}$ and $g(x) = \dfrac{-1}{1-x} + \dfrac{1}{1-x/2}$ are

$$f(x) = 1 + x + x^2 + x^3 + \cdots = \sum_{n=0}^{\infty} x^n$$

and, by replacing x with $x/2$,

$$\begin{aligned} g(x) &= -(1 + x + x^2 + x^3 + \cdots) + (1 + \tfrac{1}{2}x + \tfrac{1}{4}x^2 + \tfrac{1}{8}x^3 + \cdots) \\ &= \sum_{n=0}^{\infty} \left(-1 + \frac{1}{2^n}\right) x^n. \end{aligned}$$

The Taylor series for each of f and g has radius of convergence 1, by the Ratio Test for power series, so Theorem F69 tells us that the radius of convergence of the Taylor series for $f+g$ is at least 1. However, the Taylor series for the function $(f+g)(x) = \dfrac{1}{1-x/2}$ is

$$(f+g)(x) = 1 + \tfrac{1}{2}x + \tfrac{1}{4}x^2 + \tfrac{1}{8}x^3 + \tfrac{1}{16}x^4 + \cdots = \sum_{n=0}^{\infty} \frac{1}{2^n}x^n,$$

which has radius of convergence 2, by the Ratio Test for power series.

Worked Exercise F47

Find the Taylor series at 0 for $f(x) = \cosh x$.

Solution

We can write

$$\cosh x = \tfrac{1}{2}(e^x + e^{-x}), \quad \text{for } x \in \mathbb{R}.$$

We know that, for $x \in \mathbb{R}$,

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + \cdots$$

This is one of the basic Taylor series given in Theorem F65.

and so, by replacing x with $-x$,

$$e^{-x} = 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \cdots (-1)^n \frac{x^n}{n!} + \cdots.$$

Then, by the Sum Rule, we deduce that

$$e^x + e^{-x} = 2\left(1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots + \frac{x^{2n}}{(2n)!} + \cdots\right), \quad \text{for } x \in \mathbb{R},$$

since the odd-powered terms cancel. It then follows, by using the Multiple Rule with $\lambda = \frac{1}{2}$, that for $x \in \mathbb{R}$,

$$\cosh x = \tfrac{1}{2}(e^x + e^{-x}) = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots + \frac{x^{2n}}{(2n)!} + \cdots.$$

Exercise F66

Find the Taylor series at 0 for each of the following functions.

(a) $f(x) = \sinh x$ (b) $f(x) = \log(1-x) + \dfrac{2}{1-x}$

We now look at how we might find the Taylor series at 0 for the function $f(x) = \dfrac{1+x}{1-x}$.

We can express $\dfrac{1+x}{1-x}$ in the form $(1+x) \times \dfrac{1}{1-x}$, so using the Taylor series for $\dfrac{1}{1-x}$ we have

$$\begin{aligned}\frac{1+x}{1-x} &= (1+x) \times \frac{1}{1-x}\\ &= (1+x) \times (1 + x + x^2 + x^3 + \cdots + x^n + \cdots), \text{ for } |x| < 1.\end{aligned}$$

If we then simply multiply out these two brackets and collect together the multiples of successive powers of x, we get

$$\begin{aligned}\frac{1+x}{1-x} &= 1 \times (1 + x + x^2 + x^3 + \cdots + x^n + \cdots)\\ &\quad + x \times (1 + x + x^2 + x^3 + \cdots + x^n + \cdots)\\ &= 1 + 2x + 2x^2 + 2x^3 + \cdots + 2x^n + \cdots,\end{aligned}$$

and this power series has radius of convergence 1.

To justify multiplying together Taylor series in this way to obtain further Taylor series, we now state and prove the Product Rule for Taylor series. If you are short of time, then you may prefer to skim read the proof.

Theorem F70 Product Rule for Taylor series

Let f and g be functions that can both be represented by a Taylor series at a, and suppose that

$$f(x) = \sum_{n=0}^{\infty} a_n (x-a)^n, \quad \text{for } |x-a| < R,$$

$$g(x) = \sum_{n=0}^{\infty} b_n (x-a)^n, \quad \text{for } |x-a| < R'.$$

Then if $r = \min\{R, R'\}$ we have

$$(fg)(x) = \sum_{n=0}^{\infty} c_n (x-a)^n, \quad \text{for } |x-a| < r,$$

where

$$c_n = a_0 b_n + a_1 b_{n-1} + \cdots + a_{n-1} b_1 + a_n b_0 = \sum_{k=0}^{n} a_k b_{n-k}.$$

Note that although the expression for the coefficient c_n looks rather complicated, it is just the result of multiplying the two Taylor series term by term and summing all the resulting coefficients of $(x-a)^n$.

Proof of Theorem F70 For simplicity, we assume that $a = 0$.

Take $n \geq 2$ and put $m = \lfloor \frac{1}{2}n \rfloor$, the integer part of $\frac{1}{2}n$. Then

$$\sum_{i=0}^{n} a_i x^i \times \sum_{j=0}^{n} b_j x^j = \sum_{k=0}^{n} c_k x^k + \text{the sum of those terms } a_i b_j x^{i+j} \text{ in } \sum_{i=0}^{n} a_i x^i \times \sum_{j=0}^{n} b_j x^j \text{ with } i + j > n.$$

Thus, by the Triangle Inequality,

$$\left| \sum_{i=0}^{n} a_i x^i \times \sum_{j=0}^{n} b_j x^j - \sum_{k=0}^{n} c_k x^k \right| \leq \text{the sum of those terms } \left|a_i b_j x^{i+j}\right| \text{ in } \sum_{i=0}^{n} |a_i x^i| \times \sum_{j=0}^{n} |b_j x^j| \text{ with } i + j > n.$$

But all the latter terms are included in the expression

$$\sum_{i=0}^{n} |a_i x^i| \times \sum_{j=0}^{n} |b_j x^j| - \sum_{i=0}^{m} |a_i x^i| \times \sum_{j=0}^{m} |b_j x^j|,$$

and the other terms in this expression are all non-negative.

Note that all of the terms $\left|a_i b_j x^{i+j}\right|$ in $\sum_{i=0}^{m} |a_i x^i| \times \sum_{j=0}^{m} |b_j x^j|$ have $i + j \leq 2m \leq n$.

Hence

$$\left| \sum_{i=0}^{n} a_i x^i \times \sum_{j=0}^{n} b_j x^j - \sum_{k=0}^{n} c_k x^k \right|$$
$$\leq \sum_{i=0}^{n} |a_i x^i| \times \sum_{j=0}^{n} |b_j x^j| - \sum_{i=0}^{m} |a_i x^i| \times \sum_{j=0}^{m} |b_j x^j|. \qquad (8)$$

Now suppose that $|x| < r$, so the series $\sum_{i=0}^{\infty} |a_i x^i|$ and $\sum_{j=0}^{\infty} |b_j x^j|$ are both convergent, with sums s and t, respectively.

Here we have used the fact that a power series is absolutely convergent at each interior point of its interval of convergence.

As $n \to \infty$, the right-hand side of inequality (8) tends to $st - st = 0$, since $m \to \infty$. Thus, by the Limit Inequality Rule (Theorem D11 in Unit D2), $\sum_{k=0}^{\infty} c_k x^k$ converges, and

$$\sum_{k=0}^{\infty} c_k x^k = \sum_{i=0}^{\infty} a_i x^i \times \sum_{j=0}^{\infty} b_j x^j.$$

This completes the proof of the Product Rule. ■

Worked Exercise F48

Find the Taylor series at 0 for the function $f(x) = \dfrac{1+x}{(1-x)^2}$.

Solution

We first write

$$\frac{1+x}{(1-x)^2} = \frac{1+x}{1-x} \times \frac{1}{1-x}.$$

We know that, for $|x| < 1$,

$$\frac{1+x}{1-x} = 1 + 2x + 2x^2 + 2x^3 + \cdots + 2x^n + \cdots$$

and

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots + x^n + \cdots.$$

We discussed the first of these two series just before stating the Product Rule, and our informal argument there is justified by the proof of the Product Rule. The second series is a basic Taylor series listed in Theorem F65.

Hence, by the Product Rule,

$$\begin{aligned}\frac{1+x}{(1-x)^2} &= \frac{1+x}{1-x} \times \frac{1}{1-x}\\ &= (1 + 2x + 2x^2 + 2x^3 + \cdots + 2x^n + \cdots)\\ &\quad \times (1 + x + x^2 + x^3 + \cdots + x^n + \cdots)\\ &= 1 + (2+1)x + (2+2+1)x^2 + \cdots\\ &\quad + (2+2+\cdots+2+1)x^n + \cdots\\ &= 1 + 3x + 5x^2 + \cdots + (2n+1)x^n + \cdots,\end{aligned}$$

for $|x| < 1$.

Thus the Taylor series for f at 0 is $\displaystyle\sum_{n=0}^{\infty} (2n+1)x^n$, for $|x| < 1$.

Exercise F67

Determine the Taylor series at 0 for each of the following functions.

(a) $f(x) = (1+x)\log(1+x)$ (b) $f(x) = \dfrac{1}{(1-x)^2}$

(c) $f(x) = \dfrac{1+x}{(1-x)^3}$

Hint: In part (c), use the solution to Worked Exercise F48.

Exercise F68

Determine the Taylor series at 0 for each of the following functions. In each case, indicate the general term and state a range of validity for the series.

(a) $f(x) = \sinh x + \sin x$ (b) $f(x) = \dfrac{1}{1+2x^2}$

Hint: In part (a), use the solution to Exercise F66(a).

Exercise F69

Determine the first three non-zero terms in the Taylor series at 0 for the function $f(x) = e^x(1-x)^{-2}$, and state a range of validity for the series.

Hint: Use the solution to Exercise F67(b).

4.2 The Differentiation and Integration Rules

We have seen that the hyperbolic functions have the following Taylor series at 0:

$$\sinh x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots \quad \text{and} \quad \cosh x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots,$$

each with radius of convergence ∞. Notice that the derivative of the function $\sinh x$ is $\cosh x$, and that term-by-term differentiation of the series $x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots$ gives the series $1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots$. It looks as though we can obtain the Taylor series for the derivative of a function f simply by differentiating the Taylor series for f itself.

Our next two results show that we can differentiate or integrate the Taylor series of a function f term-by-term to obtain the Taylor series of the corresponding function f' or $\displaystyle\int f$, respectively, where we make an appropriate choice for the constant of integration.

Theorem F71 Differentiation Rule for Taylor series

The Taylor series

$$f(x) = \sum_{n=0}^{\infty} a_n(x-a)^n \quad \text{and} \quad g(x) = \sum_{n=1}^{\infty} na_n(x-a)^{n-1}$$

have the same radius of convergence, R say.

Also, $f(x)$ is differentiable on $(a-R, a+R)$, and

$$f'(x) = g(x), \quad \text{for } |x-a| < R.$$

Note that the two series in the Differentiation Rule may behave differently at the endpoints of their respective intervals of convergence. For example, the power series

$$\frac{1}{1^2}x + \frac{1}{2^2}x^2 + \frac{1}{3^2}x^3 + \cdots = \sum_{n=1}^{\infty} \frac{1}{n^2}x^n$$

and

$$\frac{1}{1^2} + \frac{1}{2^2}2x + \frac{1}{3^2}3x^2 + \cdots = 1 + \frac{1}{2}x + \frac{1}{3}x^2 + \cdots = \sum_{n=1}^{\infty} \frac{1}{n}x^{n-1}$$

both have radius of convergence 1. However, the first series converges at both ± 1, whereas the second series converges at -1 but diverges at 1.

Proof of the Differentiation Rule (optional) For simplicity, we assume that $a = 0$.

Let the series $\displaystyle\sum_{n=0}^{\infty} a_n x^n$ and $\displaystyle\sum_{n=1}^{\infty} n a_n x^{n-1}$ have radii of convergence R and R', respectively. We prove that $R' = R$.

We first show that $R' \geq R$. To prove this, suppose that $|x| < R$. Now choose r such that $|x| < r < R$, as shown in Figure 13. Then $\displaystyle\sum_{n=0}^{\infty} a_n r^n$ is convergent, so $(a_n r^n)$ is a null sequence. Thus there is a positive number K such that

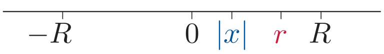


Figure 13 The relationship between r, $|x|$ and R.

$$|a_n r^n| \leq K, \quad \text{for } n = 0, 1, 2, \ldots. \tag{9}$$

Then

$$|n a_n x^{n-1}| = \frac{|n a_n r^n x^{n-1}|}{r^n} \leq \frac{K}{r} n \left(\frac{|x|}{r}\right)^{n-1}, \quad \text{for } n = 1, 2, \ldots, \tag{10}$$

by inequalities (9). Since $|x|/r < 1$, the series $\displaystyle\sum_{n=1}^{\infty} n\left(|x|/r\right)^{n-1}$ converges.

This follows from the solution to Exercise F65(a).

Therefore, by statement (10) and the Comparison Test (Theorem D30 in Unit D3), $\displaystyle\sum_{n=1}^{\infty} |n a_n x^{n-1}|$ is convergent for all $|x| < R$. This proves that $R' \geq R$.

Next we show that $R \geq R'$. To prove this, suppose that $|x| < R'$. Then $\displaystyle\sum_{n=1}^{\infty} n a_n x^{n-1}$ is absolutely convergent and

$$|a_n x^n| = |n a_n x^{n-1}| \frac{|x|}{n} \leq |x|\,|n a_n x^{n-1}|, \quad \text{for } n = 1, 2, \ldots.$$

Therefore, by the Comparison Test, $\displaystyle\sum_{n=1}^{\infty} |a_n x^n|$ is convergent for all $|x| < R'$. Thus $R \geq R'$, so we deduce that $R' = R$.

Differentiating the terms of $\sum_{n=1}^{\infty} na_n x^{n-1}$, we deduce that

$$\sum_{n=2}^{\infty} n(n-1)a_n x^{n-2} \text{ also has radius of convergence } R, \tag{11}$$

by an analogous argument to the one above.

Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$. We now use statement (11) to prove that f' exists and has the required form on $(-R, R)$.

Take $x \in (-R, R)$, and choose r such that $|x| < r < R$. Then, for all h such that $|x + h| < r$ (see Figure 14),

$$\begin{aligned}\frac{f(x+h)-f(x)}{h} - \sum_{n=1}^{\infty} na_n x^{n-1} &= \sum_{n=1}^{\infty} \frac{a_n(x+h)^n}{h} - \sum_{n=1}^{\infty} \frac{a_n x^n}{h} - \sum_{n=1}^{\infty} na_n x^{n-1} \\ &= \sum_{n=2}^{\infty} a_n \frac{(x+h)^n - x^n - nx^{n-1}h}{h}. \end{aligned} \tag{12}$$

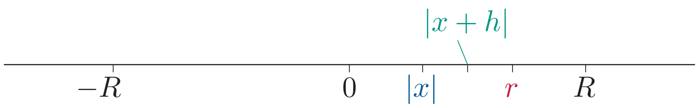


Figure 14 The point $|x + h|$

Now we apply Taylor's Theorem to the function $p(x) = x^n$ on an open interval containing x and $x + h$. We obtain

$$p(x+h) = (x+h)^n = x^n + nx^{n-1}h + \frac{1}{2!}n(n-1)c_n^{n-2}h^2,$$

where c_n lies between x and $x + h$.

By Taylor's Theorem, we have

$$p(x+h) = p(x) + p'(x)((x+h)-x) + R_1(x),$$

where $p'(x) = nx^{n-1}$ and the remainder term $R_1(x)$ depends on the second derivative $p''(x) = n(n-1)x^{n-2}$.

Then $|c_n| < r$, so

$$|(x+h)^n - x^n - nx^{n-1}h| \leq \tfrac{1}{2}n(n-1)r^{n-2}|h|^2. \tag{13}$$

By equation (12) and inequality (13), together with the Triangle Inequality, we obtain

$$\left| \frac{f(x+h)-f(x)}{h} - \sum_{n=1}^{\infty} na_n x^{n-1} \right| \leq \tfrac{1}{2}|h| \sum_{n=2}^{\infty} n(n-1)|a_n| r^{n-2}. \tag{14}$$

Since $r < R$, the series $\sum_{n=2}^{\infty} n(n-1)a_n r^{n-2}$ is absolutely convergent, by statement (11). Thus, by inequality (14) and the Limit Inequality Rule,

$$f'(x) = \lim_{h \to 0} \frac{f(x+h)-f(x)}{h} = \sum_{n=1}^{\infty} na_n x^{n-1}, \quad \text{for } |x| < R.$$

■

We can now easily obtain the Integration Rule from the Differentiation Rule.

Theorem F72 Integration Rule for Taylor series

The Taylor series

$$f(x) = \sum_{n=0}^{\infty} a_n(x-a)^n \quad \text{and} \quad F(x) = \sum_{n=0}^{\infty} \frac{a_n}{n+1}(x-a)^{n+1}$$

have the same radius of convergence, R say.

Also, if $R > 0$, then

$$\int f(x)\,dx = F(x), \quad \text{for } |x-a| < R.$$

Remarks

1. As in the Differentiation Rule, the two series in the theorem may behave differently at the endpoints of their respective intervals of convergence.
2. The final conclusion says that F is a primitive of f on $(a-R, a+R)$. It is sometimes expressed in the following way:

   $$\int \left(\sum_{n=0}^{\infty} a_n(x-a)^n \right) dx = c + \sum_{n=0}^{\infty} \frac{a_n}{n+1}(x-a)^{n+1}.$$

   We find c by putting $x = a$ into the equation.

Proof of the Integration Rule The two series have the same radius of convergence, by the Differentiation Rule applied to F.

By the same rule, $F' = f$, so F is a primitive of f on $(a-R, a+R)$. ■

Worked Exercise F49

Find the Taylor series at 0 for $f(x) = \tan^{-1} x$.

Solution

We know that

$$\frac{d}{dx}\tan^{-1} x = \frac{1}{1+x^2}, \quad \text{for } x \in \mathbb{R}.$$

We also know that, for $|x| < 1$,

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \cdots + (-1)^n x^{2n} + \cdots.$$

You can see this by replacing x by $-x^2$ in the basic series for $1/(1-x)$.

Hence, by the Integration Rule,

$$\tan^{-1} x = c + x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots + (-1)^n \frac{x^{2n+1}}{2n+1},$$

for $|x| < 1$, where c is a constant.

Substituting $x = 0$ into this equation, we find that $c = \tan^{-1} 0 = 0$. Hence

$$\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots + \frac{(-1)^n x^{2n+1}}{2n+1} + \cdots,$$

for $|x| < 1$.

In fact the above formula also holds for $x = 1$ and $x = -1$, but the proof of these facts is somewhat more complicated and we do not give it here.

Exercise F70

Find the Taylor series at 0 for:

(a) $f(x) = (1-x)^{-3}$ (b) $f(x) = \tanh^{-1} x$.

Hint: You may find the solution to Exercise F67(b) helpful in part (a).

Exercise F71

Determine the Taylor series at 0 for the function $f(x) = e^{-x^2}$. Deduce that

$$\int_0^1 e^{-x^2}\,dx = 1 - \frac{1}{3} + \frac{1}{10} - \cdots + \frac{(-1)^n}{(2n+1)n!} + \cdots.$$

Exercise F72

Use the Taylor series for $1/(1+x)$ at 0 to determine the Taylor series for the function $g(x) = \log(1+x)$ at the same point, and state a range of validity for this series.

In Exercise F72 you established that the Taylor series at 0 for the function $f(x) = \log(1+x)$ is valid on the interval $(-1, 1)$. In Section 2 we proved that this Taylor series is valid on $[0, 1]$, so by combining these results we see that the full range of validity for the Taylor series at 0 for the function $f(x) = \log(1+x)$ is the interval $(-1, 1]$, as stated in Theorem F65(e).

4.3 The General Binomial Theorem and the Uniqueness Theorem

In Subsection 3.4 of Unit D1 *Numbers* you met the Binomial Theorem, which states that, for each positive integer n,

$$(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k,$$

where

$$\binom{n}{0} = 1 \quad \text{and} \quad \binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\cdots(n-k+1)}{k!}, \quad \text{for } k \in \mathbb{N}.$$

This gives the Taylor series for $(1+x)^n$ and is valid for all $x \in \mathbb{R}$. In fact, a similar result known as the *General Binomial Theorem* holds for more general powers of $(1+x)$, but with a restriction on the values of x for which it is valid.

Theorem F73 General Binomial Theorem

For $\alpha \in \mathbb{R}$,

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n, \quad \text{for } |x| < 1,$$

where

$$\binom{\alpha}{0} = 1 \quad \text{and} \quad \binom{\alpha}{n} = \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}, \quad \text{for } n \in \mathbb{N}.$$

Remarks

1. The Binomial Theorem is usually stated as a sum of powers of x^k. We have stated the General Binomial Theorem as a sum of powers of x^n in order to match the Taylor series in the rest of the unit. Note that this means that the role of n in the generalised binomial coefficient is the same as the role of k in the normal binomial coefficient.
2. The coefficient $\binom{\alpha}{n}$ is known as a **generalised binomial coefficient**. An example of calculating a generalised binomial coefficient is

$$\begin{aligned}\binom{-\frac{1}{2}}{n} &= \frac{\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)\left(-\frac{5}{2}\right)\cdots\left(-\frac{1}{2}-n+1\right)}{n!} \\ &= (-1)^n \frac{1\cdot 3\cdot 5\cdot \cdots \cdot (2n-1)}{2^n n!}.\end{aligned}$$

We now give the proof of the General Binomial Theorem. If you are short of time, then you may wish to skim read this proof.

Proof of the General Binomial Theorem

Let

$$f(x) = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n \quad \text{and} \quad g(x) = f(x)(1+x)^{-\alpha}, \qquad \text{for } |x| < 1.$$

We first note that the series $f(x) = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n$ converges for $|x| < 1$, as proved in Exercise F64. We want to prove that $g(x) = 1$ and hence that $f(x) = (1+x)^{\alpha}$, for all x with $|x| < 1$.

Using the Product Rule for differentiation, we differentiate the expression for g to obtain

$$\begin{aligned} g'(x) &= f'(x)(1+x)^{-\alpha} - \alpha f(x)(1+x)^{-\alpha-1} \\ &= ((1+x)f'(x) - \alpha f(x))(1+x)^{-\alpha-1}. \end{aligned} \tag{15}$$

Now, using the Differentiation Rule,

$$\begin{aligned} (1+x)f'(x) - \alpha f(x) &= (1+x)\sum_{n=1}^{\infty} n\binom{\alpha}{n} x^{n-1} - \alpha \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n \\ &= \sum_{n=1}^{\infty} n\binom{\alpha}{n} x^{n-1} + \sum_{n=1}^{\infty} n\binom{\alpha}{n} x^n - \alpha\sum_{n=0}^{\infty} \binom{\alpha}{n} x^n \\ &= \sum_{n=0}^{\infty} (n+1)\binom{\alpha}{n+1} x^n + \sum_{n=0}^{\infty} (n-\alpha)\binom{\alpha}{n} x^n, \end{aligned}$$

since

$$\sum_{n=1}^{\infty} n\binom{\alpha}{n} x^{n-1} = \sum_{n=0}^{\infty} (n+1)\binom{\alpha}{n+1} x^n,$$

and

$$\sum_{n=1}^{\infty} n\binom{\alpha}{n} x^n = \sum_{n=0}^{\infty} n\binom{\alpha}{n} x^n.$$

Hence

$$(1+x)f'(x) - \alpha f(x) = \sum_{n=0}^{\infty} \left((n+1)\binom{\alpha}{n+1} + (n-\alpha)\binom{\alpha}{n} \right) x^n. \tag{16}$$

We now use algebraic manipulation and the definition of the generalised binomial coefficient to simplfy the expression in square brackets in equation (16).

$$\begin{aligned}(n+1)\binom{\alpha}{n+1}+(n-\alpha)\binom{\alpha}{n} &= (n+1)\frac{\alpha(\alpha-1)\cdots(\alpha-n+1)(\alpha-n)}{(n+1)!} \\ &\quad +(n-\alpha)\binom{\alpha}{n} \\ &= \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}(\alpha-n) \\ &\quad +(n-\alpha)\binom{\alpha}{n} \\ &= \binom{\alpha}{n}(\alpha-n)+(n-\alpha)\binom{\alpha}{n} \\ &= 0.\end{aligned}$$

Working backwards, it follows from equation (16) that $(1+x)f'(x)-\alpha f(x)=0$, and then from equation (15) that $g'(x)=0$.

So, $g(x)$ is a constant. Hence

$$\begin{aligned}g(x)=g(0)&=f(0)(1+0)^{-\alpha} \\ &=f(0)=\binom{\alpha}{0}=1,\end{aligned}$$

as required. ■

Worked Exercise F50

Use the General Binomial Theorem to find the first three non-zero terms in the Taylor series at 0 for the function $f(x)=(1+2x)^{-6}$.

Solution

By the General Binomial Theorem,

$$(1+2x)^{-6}=\sum_{n=0}^{\infty}\binom{-6}{n}(2x)^n,\quad \text{for } |2x|<1,$$

where

$$\binom{-6}{0}=1 \quad\text{and}\quad \binom{-6}{n}=\frac{(-6)(-7)\cdots(-6-n+1)}{n!},\quad \text{for } n\in\mathbb{N}.$$

Hence

$$(1+2x)^{-6}=1+\frac{(-6)}{1}2x+\frac{(-6)(-7)}{2}(2x)^2+\cdots,\quad \text{for } |2x|<1.$$

So

$$(1+2x)^{-6}=1-12x+84x^2-\cdots,\quad \text{for } |x|<\tfrac{1}{2}.$$

Exercise F73

Use the General Binomial Theorem to find the first three non-zero terms in the Taylor series at 0 for the function $f(x) = (1 + 4x)^{-1/3}$.

We now have a variety of techniques for finding Taylor series:

- Taylor's Theorem
- the Combination Rules
- the Product Rule
- the Differentiation and Integration Rules
- the General Binomial Theorem.

But how do we know that these different techniques will always give us the same expression for the Taylor series of a given function? To end this section we prove a result which states that there is only one Taylor series for a function f at a given point a. Thus any valid method gives the same Taylor coefficients.

Theorem F74 Uniqueness Theorem for Taylor series

If

$$\sum_{n=0}^{\infty} a_n(x-a)^n = \sum_{n=0}^{\infty} b_n(x-a)^n, \quad \text{for } |x-a| < R,$$

then $a_n = b_n$, for $n = 0, 1, 2, \ldots$.

Proof Let

$$f(x) = \sum_{n=0}^{\infty} a_n(x-a)^n \quad \text{and} \quad g(x) = \sum_{n=0}^{\infty} b_n(x-a)^n, \quad \text{for } |x-a| < R.$$

If we differentiate both equations n times using the Differentiation Rule, and put $x = a$, then we obtain

$$f^{(n)}(a) = (n!)a_n \quad \text{and} \quad g^{(n)}(a) = (n!)b_n.$$

Since $f(x) = g(x)$ for $|x-a| < R$, it follows that $f^{(n)}(a) = g^{(n)}(a)$. Hence $a_n = b_n$, for all $n = 0, 1, 2, \ldots$. ■

5 Numerical estimates for π

One of the problems that has fascinated mathematicians for thousands of years has been how to determine accurately various important irrational numbers such as $\sqrt{2}$, π and e. In this section you will see the role of Taylor series in the numerical estimation of π, and meet an ingenious proof that π is irrational.

Historical estimates for π

The use of the symbol π to denote the ratio of the circumference of a circle to its diameter is relatively new, being first introduced by William Jones in 1706. Its use was popularised by Leonhard Euler who employed it in his *Introductio in Analysin Infinitorum* of 1748.

Early estimates for π include the following:

Mesopotamia (c.2000 BCE) In 1936 a Babylonian clay tablet was excavated which gives the ratio of the perimeter of a regular hexagon to the circumference of the circumscribed circle as $\frac{57}{60} + \frac{36}{(60)^2}$ (the Babylonians used the sexagesimal (base 60) system rather than the decimal system). This corresponds to a value for π of 3.125.

Egypt (c.1900 BCE) The Egyptians found by experience that they could approximate the area of a circle with diameter d by reducing d by one-ninth and squaring it. This corresponds to a value for π of $\frac{256}{81} \approx 3.1605$.

Archimedes (c.250 BCE) Archimedes (c.287–c.212 BCE) considered hexagons inside and outside a circle and compared their perimeters with the circumference of the circle. This gave him a value for π between 3 and 3.464. He then replaced the hexagon with a 12-sided polygon and recalculated the lengths. He continued by progressively doubling the sides of the polygon until he reached a polygon of 96 sides. This gave him a value for π between $3\frac{10}{71}$ and $3\frac{1}{7}$, or $3.14084 < \pi < 3.14286$, which is correct to two decimal places. Archimedes' method was described in some detail in Subsection 5.2 of Unit D2.

China (c.100–500) Zhang Heng (78–139) considered the ratio of the area of a square to the area of a circle and the ratio of the volume of a cube to the volume of a sphere, and was led to an approximation for π of $\sqrt{10}\,(\approx 3.162)$. He also calculated π as $\frac{736}{232}\,(\approx 3.1724)$. Liu Hui (c.220–c.280) used a polygonal method similar to that of Archimedes. He doubled the sides of a regular polygon until he reached a polygon with 3072 sides, from which he calculated a value for π of 3.14159. Zu Chongzhi (429–500) deduced that $3.1415926 < \pi < 3.1415927$, which was the most accurate approximation for π for almost a millennium. It is not known how he obtained this result but it is possible that he considered polygons with $24\,576\,(= 2^{13} \times 3)$ sides.

India (499) Aryabhata (476–550) in his *Aryabhatiya* included the following statement: 'Add 4 to 100, multiply by 8 and add 62 000. The result is approximately the circumference of a circle of which the diameter is 20 000.' This gives an approximate value for π of 3.1416. Although Aryabhata does not say how this value was found, it is likely that it was done by the method of polygons, using polygons with 384 sides.

With the development of calculus in the seventeenth century, new formulas for estimating π were discovered, including Wallis' Formula which you met in Subsection 3.2 of Unit F3 *Integration*:

$$\frac{\pi}{2} = \frac{2}{1}\cdot\frac{2}{3}\cdot\frac{4}{3}\cdot\frac{4}{5}\cdot\frac{6}{5}\cdot\frac{6}{7}\cdot\cdots\cdot\frac{2n}{2n-1}\cdot\frac{2n}{2n+1}\cdots.$$

5.1 Tangent formulas

In Worked Exercise F49 you saw that the Taylor series at 0 for the function $\tan^{-1}$ is

$$\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots, \quad \text{for } x \in [-1, 1].$$

In particular, with $x = 1$,

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots.$$

The first of these two series is known as *Leibniz's series* and the second as *Gregory's series*, although these names are often interchanged.

The second series is not very useful for calculating π, as its successive partial sums converge far too slowly. The smaller the value of x, the faster the Taylor series for $\tan^{-1} x$ converges, so fewer terms are needed to calculate its sum to a given accuracy.

To obtain series that are more effective for calculating π, we can use an addition formula for $\tan^{-1}$. To derive this formula, first recall the addition formula for tan (given in the module Handbook):

$$\tan(a+b) = \frac{\tan a + \tan b}{1 - \tan a \tan b}.$$

If we now put $x = \tan a$ and $y = \tan b$, we have

$$\tan(a+b) = \frac{x+y}{1-xy},$$

so it follows that

$$\tan^{-1} x + \tan^{-1} y = \tan^{-1}\left(\frac{x+y}{1-xy}\right), \quad \text{for } x, y \in \mathbb{R}, \tag{17}$$

provided that $\tan^{-1} x + \tan^{-1} y$ lies in $(-\pi/2, \pi/2)$, the image set of $\tan^{-1}$.

For example, applying formula (17) with $x = \frac{1}{2}$ and $y = \frac{1}{3}$, we obtain

$$\tan^{-1}\left(\tfrac{1}{2}\right) + \tan^{-1}\left(\tfrac{1}{3}\right) = \tan^{-1} 1 = \pi/4.$$

Using the addition formula (17) with small values of x combined with the Taylor series for $\tan^{-1}$ at 0 gives efficient ways of calculating π. For example, repeated application of the formula gives the following (we omit the details):

$$4\tan^{-1}\left(\tfrac{1}{5}\right) - \tan^{-1}\left(\tfrac{1}{239}\right) = \pi/4,$$
$$6\tan^{-1}\left(\tfrac{1}{8}\right) + 2\tan^{-1}\left(\tfrac{1}{57}\right) + \tan^{-1}\left(\tfrac{1}{239}\right) = \pi/4.$$

The first of these formulas is called Machin's Formula. John Machin (1680–1751) used the formula to calculate the first 100 decimal places of π, and in 1974 such formulas were used to calculate π to a million decimal places. More recently, highly ingenious methods (based on techniques due to Gauss for evaluating integrals approximately) have been used to calculate π correct to many billions of decimal places.

For your interest, we now list the first 1000 decimal places of π.

3.1415926535 8979323846 2643383279 5028841971 6939937510
5820974944 5923078164 0628620899 8628034825 3421170679
8214808651 3282306647 0938446095 5058223172 5359408128
4811174502 8410270193 8521105559 6446229489 5493038196
4428810975 6659334461 2847564823 3786783165 2712019091
4564856692 3460348610 4543266482 1339360726 0249141273
7245870066 0631558817 4881520920 9628292540 9171536436
7892590360 0113305305 4882046652 1384146951 9415116094
3305727036 5759591953 0921861173 8193261179 3105118548
0744623799 6274956735 1885752724 8912279381 8301194912
9833673362 4406566430 8602139494 6395224737 1907021798
6094370277 0539217176 2931767523 8467481846 7669405132
0005681271 4526356082 7785771342 7577896091 7363717872
1468440901 2249534301 4654958537 1050792279 6892589235
4201995611 2129021960 8640344181 5981362977 4771309960
5187072113 4999999837 2978049951 0597317328 1609631859
5024459455 3469083026 4252230825 3344685035 2619311881
7101000313 7838752886 5875332083 8142061717 7669147303
5982534904 2875546873 1159562863 8823537875 9375195778
1857780532 1712268066 1300192787 6611195909 2164201989

There are several mnemonics that can be used to recall the first few digits of π, with the word lengths giving the successive digits of π. For example, one such mnemonic is:

May	I	have	a	large	container	of	coffee?
3.	1	4	1	5	9	2	6

5.2 Proof that π is irrational (optional)

We finish the unit with an interesting proof which uses several ideas that you have met in earlier analysis units. You began your study of analysis in this module with Unit D1 where you first considered rational and irrational numbers, which together make up the real numbers whose properties lie at the foundation of analysis. Now we come full circle and use some of the tools of analysis that you have learned to prove that π is irrational.

The first proof that π is irrational was given by Johann Heinrich Lambert in 1766, but the elegant, shorter proof that we give here was found by Ivan Niven in 1947.

Theorem F75

The number π is irrational.

Proof We prove that π^2 is irrational, from which it follows that π is irrational. The proof is by contradiction and the method is rather unusual, so we begin by outlining the two major steps.

First we show that if f is any polynomial function such that

$$0 < f''(x) + \pi^2 f(x) < 1, \quad \text{for } 0 < x < 1, \tag{18}$$

then

$$0 < f(0) + f(1) < \frac{1}{\pi}. \tag{19}$$

Next we show that if $\pi^2 = a/b$, for $a, b \in \mathbb{N}$, then there is a polynomial function f such that statement (18) is true but inequalities (19) do not hold. This contradiction shows that π^2 must be irrational.

Let f be a polynomial function satisfying statement (18), and put

$$g(x) = f'(x)\sin \pi x - \pi f(x) \cos \pi x.$$

Then

$$\begin{aligned} g'(x) &= f''(x)\sin \pi x + f'(x)\pi \cos \pi x - \pi f'(x)\cos \pi x + \pi^2 f(x) \sin \pi x \\ &= \big(f''(x) + \pi^2 f(x)\big) \sin \pi x. \end{aligned}$$

By the Mean Value Theorem (Theorem F37 in Unit F2), there exists $c \in (0, 1)$ such that

$$g(1) - g(0) = g'(c) = \big(f''(c) + \pi^2 f(c)\big) \sin \pi c.$$

Hence $0 < g(1) - g(0) < 1$, by statement (18) and the fact that $0 < \sin \pi c \leq 1$. But

$$g(1) - g(0) = \pi(f(0) + f(1)),$$

by the definition of g, so statement (19) follows.

Now suppose that $\pi^2 = a/b$, where $a, b \in \mathbb{N}$. Take $N \in \mathbb{N}$ so large that $\pi^2 a^N/N! < 1$,

This is possible because $(a^n/n!)$ is a basic null sequence: see Theorem D7 in Unit D2.

and put

$$p(x) = \frac{1}{N!}\,x^N(1-x)^N \tag{20}$$
$$= \frac{1}{N!}\left(c_N x^N + c_{N+1}x^{N+1} + \cdots + c_{2N}x^{2N}\right),$$

where the coefficients c_k are integers, for $N \leq k \leq 2N$. Then we have

$$0 < p(x) < 1/N!, \quad \text{for } 0 < x < 1, \tag{21}$$

by equation (20). Also, for $k = 0, 1, \ldots,$

$$p^{(k)}(0) = \begin{cases} 0, & 0 \leq k < N,\ k > 2N, \\ \dfrac{c_k k!}{N!}, & N \leq k \leq 2N, \end{cases}$$

so $p^{(k)}(0)$ is an integer. Hence $p^{(k)}(1)$ is also an integer, by the symmetry of the function p under the change of variable $x' = 1 - x$.

Now consider the polynomial function

$$f(x) = a^N p(x) - a^{N-1}b\,p^{(2)}(x) + \cdots + (-1)^N b^N p^{(2N)}(x),$$

which has degree $2N$. Then $f(0)$ and $f(1)$ are both integers, so statement (19) is false, since we know that $\pi > 1$. Finally,

$$f''(x) = a^N p^{(2)}(x) - a^{N-1}b\,p^{(4)}(x) + \cdots + (-1)^{N-1}ab^{N-1}p^{(2N)}(x),$$

since $p^{(2N+2)} = 0$. Hence

$$f''(x) + \pi^2 f(x) = f''(x) + \frac{a}{b}\,f(x) = \frac{a}{b}a^N p(x) = \pi^2 a^N p(x),$$

by telescopic cancellation. Thus, by statement (21) and our choice of N,

$$0 < f''(x) + \pi^2 f(x) < \pi^2 a^N/N! < 1, \quad \text{for } 0 < x < 1,$$

so statement (18) does hold. This completes the proof. ■

Summary

In this unit you have seen that many functions can be approximated by Taylor polynomials: for a function f that is n-times differentiable on an open interval containing the point a, the Taylor polynomial of degree n at a for f is the polynomial

$$T_n(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

You also saw how to estimate the accuracy of this approximation using Taylor's Theorem and an upper bound for a remainder term $R_n(x)$ that depends on the $(n+1)$th derivative of f. For a function f that has derivatives of all orders on an open interval, you learned that at points where the remainder term tends to zero as $n \to \infty$, f can be represented by a convergent power series known as the Taylor series for f.

You went on to study how functions can be defined by means of power series, and how the Ratio Test can be used to find their radius of convergence. You saw how to manipulate Taylor series to obtain new series from old by using the Combination Rules, the Product Rule and the Integration and Differentiation Rules. You also met the General Binomial Theorem, which gives the Taylor series at 0 for the function $(1+x)^\alpha$ for $\alpha \in \mathbb{R}$. Finally, you looked at how the Taylor series at 0 for $\tan^{-1} x$ can be used to obtain estimates for the value of π.

Learning outcomes

After working through this unit, you should be able to:

- calculate the *Taylor polynomial* $T_n(x)$ at a given point a of a given function f
- appreciate that in many cases $T_n(x)$ gives a good approximation to $f(x)$ for x near the point a
- state and use Taylor's Theorem
- appreciate that a sequence of Taylor polynomials may or may not converge at a given point to the value of the function at that point
- state and use certain basic Taylor series
- state the Radius of Convergence Theorem
- determine the *radius of convergence* and the *interval of convergence* of certain power series
- state and use the Combination Rules, Product Rule, Differentiation Rule and Integration Rule for Taylor series
- state and use the General Binomial Theorem
- understand and use the Uniqueness Theorem for Taylor series.

Table of standard Taylor series

Function	Taylor series	Domain
$\dfrac{1}{1-x}$	$1+x+x^2+x^3+\cdots=\displaystyle\sum_{n=0}^{\infty}x^n$	$\lvert x\rvert<1$
$\log(1+x)$	$x-\dfrac{x^2}{2}+\dfrac{x^3}{3}-\dfrac{x^4}{4}+\cdots=\displaystyle\sum_{n=1}^{\infty}\frac{(-1)^{n+1}x^n}{n}$	$-1<x\leq 1$
e^x	$1+x+\dfrac{x^2}{2!}+\dfrac{x^3}{3!}+\cdots=\displaystyle\sum_{n=0}^{\infty}\frac{x^n}{n!}$	$x\in\mathbb{R}$
$\sin x$	$x-\dfrac{x^3}{3!}+\dfrac{x^5}{5!}-\dfrac{x^7}{7!}+\cdots=\displaystyle\sum_{n=0}^{\infty}\frac{(-1)^n x^{2n+1}}{(2n+1)!}$	$x\in\mathbb{R}$
$\cos x$	$1-\dfrac{x^2}{2!}+\dfrac{x^4}{4!}-\dfrac{x^6}{6!}+\cdots=\displaystyle\sum_{n=0}^{\infty}\frac{(-1)^n x^{2n}}{(2n)!}$	$x\in\mathbb{R}$
$(1+x)^\alpha$	$1+\alpha x+\dfrac{\alpha(\alpha-1)}{2!}x^2+\dfrac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3+\cdots=\displaystyle\sum_{n=0}^{\infty}\binom{\alpha}{n}x^n$	$\lvert x\rvert<1,\ \alpha\in\mathbb{R}$
$\sinh x$	$x+\dfrac{x^3}{3!}+\dfrac{x^5}{5!}+\dfrac{x^7}{7!}+\cdots=\displaystyle\sum_{n=0}^{\infty}\frac{x^{2n+1}}{(2n+1)!}$	$x\in\mathbb{R}$
$\cosh x$	$1+\dfrac{x^2}{2!}+\dfrac{x^4}{4!}+\dfrac{x^6}{6!}+\cdots=\displaystyle\sum_{n=0}^{\infty}\frac{x^{2n}}{(2n)!}$	$x\in\mathbb{R}$
$\tan^{-1}x$	$x-\dfrac{x^3}{3}+\dfrac{x^5}{5}-\dfrac{x^7}{7}+\cdots=\displaystyle\sum_{n=0}^{\infty}\frac{(-1)^n x^{2n+1}}{2n+1}$	$\lvert x\rvert\leq 1$

Solutions to exercises

Solution to Exercise F55

The tangent approximation to f at a is

$$f(x) \approx f(a) + f'(a)(x-a).$$

(a) We have

$$\begin{aligned} f(x) &= e^x, & f(2) &= e^2 \\ f'(x) &= e^x, & f'(2) &= e^2. \end{aligned}$$

Hence the tangent approximation to f at 2 is

$$e^x \approx e^2 + e^2(x-2) = e^2(x-1).$$

(b) We have

$$\begin{aligned} f(x) &= \cos x, & f(0) &= 1 \\ f'(x) &= -\sin x, & f'(0) &= 0. \end{aligned}$$

Hence the tangent approximation to f at 0 is

$$\cos x \approx 1 + 0(x-0) = 1.$$

Solution to Exercise F56

(a) We have

$$\begin{aligned} f(x) &= e^x, & f(2) &= e^2 \\ f'(x) &= e^x, & f'(2) &= e^2 \\ f''(x) &= e^x, & f''(2) &= e^2 \\ f^{(3)}(x) &= e^x, & f^{(3)}(2) &= e^2. \end{aligned}$$

Hence

$$T_1(x) = f(2) + f'(2)(x-2) = e^2 + e^2(x-2)$$

$$\begin{aligned} T_2(x) &= T_1(x) + \frac{f''(2)}{2!}(x-2)^2 \\ &= e^2 + e^2(x-2) + \tfrac{1}{2}e^2(x-2)^2 \end{aligned}$$

$$\begin{aligned} T_3(x) &= T_2(x) + \frac{f^{(3)}(2)}{3!}(x-2)^3 \\ &= e^2 + e^2(x-2) + \tfrac{1}{2}e^2(x-2)^2 \\ &\quad + \tfrac{1}{6}e^2(x-2)^3. \end{aligned}$$

(b) We have

$$\begin{aligned} f(x) &= \cos x, & f(0) &= 1 \\ f'(x) &= -\sin x, & f'(0) &= 0 \\ f''(x) &= -\cos x, & f''(0) &= -1 \\ f^{(3)}(x) &= \sin x, & f^{(3)}(0) &= 0. \end{aligned}$$

Hence

$$T_1(x) = f(0) + f'(0)x = 1$$

$$T_2(x) = T_1(x) + \frac{f''(0)}{2!}x^2 = 1 - \tfrac{1}{2}x^2$$

$$T_3(x) = T_2(x) + \frac{f^{(3)}(0)}{3!}x^3 = 1 - \tfrac{1}{2}x^2.$$

Solution to Exercise F57

The Taylor polynomial of degree 4 for f at a is

$$\begin{aligned} T_4(x) = f(a) + f'(a)(x-a) &+ \frac{f''(a)}{2!}(x-a)^2 \\ + \frac{f^{(3)}(a)}{3!}(x-a)^3 &+ \frac{f^{(4)}(a)}{4!}(x-a)^4. \end{aligned}$$

(a) We have

$$\begin{aligned} f(x) &= \log(1+x), & f(0) &= 0 \\ f'(x) &= 1/(1+x), & f'(0) &= 1 \\ f''(x) &= -1/(1+x)^2, & f''(0) &= -1 \\ f^{(3)}(x) &= 2/(1+x)^3, & f^{(3)}(0) &= 2 \\ f^{(4)}(x) &= -6/(1+x)^4, & f^{(4)}(0) &= -6. \end{aligned}$$

Hence

$$T_4(x) = x - \tfrac{1}{2}x^2 + \tfrac{1}{3}x^3 - \tfrac{1}{4}x^4.$$

(b) We have

$$\begin{aligned} f(x) &= \sin x, & f(\pi/4) &= 1/\sqrt{2} \\ f'(x) &= \cos x, & f'(\pi/4) &= 1/\sqrt{2} \\ f''(x) &= -\sin x, & f''(\pi/4) &= -1/\sqrt{2} \\ f^{(3)}(x) &= -\cos x & f^{(3)}(\pi/4) &= -1/\sqrt{2} \\ f^{(4)}(x) &= \sin x, & f^{(4)}(\pi/4) &= 1/\sqrt{2}. \end{aligned}$$

Hence

$$\begin{aligned} T_4(x) = &\frac{1}{\sqrt{2}}\left(1 + \left(x - \frac{\pi}{4}\right) - \frac{1}{2}\left(x - \frac{\pi}{4}\right)^2\right. \\ &\left. - \frac{1}{6}\left(x - \frac{\pi}{4}\right)^3 + \frac{1}{24}\left(x - \frac{\pi}{4}\right)^4\right). \end{aligned}$$

(c) Since $f(x) = 1 + \frac{1}{2}x - \frac{1}{2}x^2 - \frac{1}{6}x^3 + \frac{1}{4}x^4$, we have $f(0) = 1$ and

$$\begin{aligned} f'(x) &= \tfrac{1}{2} - x - \tfrac{1}{2}x^2 + x^3, & f'(0) &= \tfrac{1}{2} \\ f''(x) &= -1 - x + 3x^2, & f''(0) &= -1 \\ f^{(3)}(x) &= -1 + 6x, & f^{(3)}(0) &= -1 \\ f^{(4)}(x) &= 6, & f^{(4)}(0) &= 6. \end{aligned}$$

Hence

$$T_4(x) = 1 + \tfrac{1}{2}x - \tfrac{1}{2}x^2 - \tfrac{1}{6}x^3 + \tfrac{1}{4}x^4.$$

(Note that $T_4(x) = f(x)$, as you might expect since f is a polynomial of degree 4.)

Solution to Exercise F58

From Worked Exercise F41(a), we have

$$T_3(x) = x - \tfrac{1}{6}x^3,$$

so

$$T_3(0.1) = 0.1 - 0.001/6 = 0.099\,8\overline{3}.$$

Since

$$\sin(0.1) = 0.099\,833\,416\ldots,$$

we have

$$\begin{aligned} |\sin(0.1) - T_3(0.1)| &= 0.099\,833\,416\ldots - 0.099\,8\overline{3} \\ &\le 0.099\,833\,417 - 0.099\,833\,333 \\ &= 0.000\,000\,084 < 1 \times 10^{-7}, \end{aligned}$$

as required.

Solution to Exercise F59

(a) We have

$$f(x) = e^x, \quad f(0) = 1,$$

and in general, for each positive integer k,

$$f^{(k)}(x) = e^x, \quad f^{(k)}(0) = 1.$$

Hence

$$T_n(x) = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}.$$

(b) We have

$$\begin{aligned} f(x) &= \log(1+x), & f(0) &= 0 \\ f'(x) &= 1/(1+x), & f'(0) &= 1 \\ f''(x) &= -1/(1+x)^2, & f''(0) &= -1 \\ f^{(3)}(x) &= 2/(1+x)^3, & f^{(3)}(0) &= 2 \end{aligned}$$

and in general, for each positive integer k,

$$f^{(k)}(x) = \frac{(-1)^{k+1}(k-1)!}{(1+x)^k},$$

so that

$$f^{(k)}(0) = (-1)^{k+1}(k-1)!.$$

Hence

$$T_n(x) = x - \tfrac{1}{2}x^2 + \tfrac{1}{3}x^3 - \cdots + (-1)^{n+1}\frac{x^n}{n}.$$

(c) We have

$$\begin{aligned} f(x) &= \cos x, & f(0) &= 1 \\ f'(x) &= -\sin x, & f'(0) &= 0 \\ f''(x) &= -\cos x, & f''(0) &= -1 \\ f^{(3)}(x) &= \sin x, & f^{(3)}(0) &= 0 \\ f^{(4)}(x) &= \cos x, & f^{(4)}(0) &= 1 \end{aligned}$$

and in general, for $k = 0, 1, 2, \ldots$,

$$f^{(2k)}(0) = (-1)^k \quad \text{and} \quad f^{(2k+1)}(0) = 0.$$

Hence, for $m = 0, 1, 2, \ldots$,

$$\begin{aligned} T_{2m}(x) &= \sum_{k=0}^{m} (-1)^k \frac{x^{2k}}{(2k)!} \\ &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots + (-1)^m \frac{x^{2m}}{(2m)!}, \end{aligned}$$

and $T_{2m+1}(x) = T_{2m}(x)$.

So if n takes either of the values $2m$ or $2m+1$, then we have

$$T_n(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots + (-1)^m \frac{x^{2m}}{(2m)!}.$$

Solution to Exercise F60

From Exercise F59(c), the nth Taylor polynomial at 0 for f is

$$T_n(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots + (-1)^m \frac{x^{2m}}{(2m)!},$$

where n is either $2m$ or $2m+1$. Hence, by Taylor's Theorem with $a = 0$, $f(x) = \cos x$ and $n = 3$,

$$\begin{aligned} \cos x &= T_3(x) + R_3(x) \\ &= 1 - \tfrac{1}{2}x^2 + R_3(x), \end{aligned}$$

where

$$R_3(x) = \frac{f^{(4)}(c)}{4!}x^4 = \frac{\cos c}{4!}x^4,$$

for some c between 0 and x, as required.

Solution to Exercise F61

(a) By the solution to Exercise F59(b), we have

$$T_2(x) = x - \tfrac{1}{2}x^2.$$

(b) We use Strategy F11 with $a = 0$, $x = 0.02$ and $n = 2$.

1. First, $f^{(3)}(x) = \dfrac{2}{(1+x)^3}$; see Exercise F59(b).
2. Thus
$$|f^{(3)}(c)| = \frac{2}{(1+c)^3} \leq 2, \quad \text{for } c \in [0, 0.2],$$
so we can take $M = 2$.
3. Hence
$$\begin{aligned} |\log(1.02) - T_2(0.02)| &= |R_2(0.02)| \\ &\leq \frac{M}{(2+1)!}|x-a|^{2+1} \\ &= \frac{2}{3!} \times |0.02 - 0|^3 \\ &= 0.000\,002\overline{6} \\ &< 3 \times 10^{-6}, \end{aligned}$$
as required.

(c) By part (a),

$$T_2(0.02) = 0.02 - \tfrac{1}{2}(0.02)^2 = 0.0198.$$

By part (b),

$$|\log(1.02) - T_2(0.02)| < 0.000\,003.$$

Hence

$$0.019\,797 < \log(1.02) < 0.019\,803,$$

so

$$\log(1.02) = 0.0198 \text{ (to 4 d.p.)}.$$

Solution to Exercise F62

(a) Using the derivatives of f found in the solution to Exercise F59(c), we obtain

$$f(\pi) = -1, \quad f'(\pi) = 0, \quad f''(\pi) = 1,$$
$$f^{(3)}(\pi) = 0, \quad f^{(4)}(\pi) = -1.$$

Hence

$$T_4(x) = -1 + \tfrac{1}{2}(x-\pi)^2 - \tfrac{1}{24}(x-\pi)^4.$$

(b) We use Strategy F12 with $I = [3\pi/4, 5\pi/4]$, $a = \pi$, $r = \pi/4$ and $n = 4$.

1. First, $f^{(5)}(x) = -\sin x$.
2. Thus
$$|f^{(5)}(c)| \leq 1, \quad \text{for } c \in [3\pi/4, 5\pi/4],$$
so we can take $M = 1$.
3. Hence
$$\begin{aligned} |R_4(x)| &\leq \frac{M}{(4+1)!}r^{4+1} \\ &= \frac{1}{5!}\left(\frac{\pi}{4}\right)^5 \\ &= 0.002\,49\ldots \\ &< 3 \times 10^{-3}, \quad \text{for } x \in [3\pi/4, 5\pi/4]. \end{aligned}$$

Thus $T_4(x)$ approximates $f(x)$ with an error less than 3×10^{-3} on $[3\pi/4, 5\pi/4]$.

Solution to Exercise F63

(a) Here $a_n = 2^n + 4^n$, for $n = 0, 1, 2, \ldots$, so

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{2^{n+1} + 4^{n+1}}{2^n + 4^n}.$$

Dividing the numerator and denominator by 4^n gives

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{2(1/2)^n + 4}{(1/2)^n + 1} \to 4 \text{ as } n \to \infty.$$

Hence, by the Ratio Test, the radius of convergence is $R = \frac{1}{4}$.

(b) Here $a_n = (n!)^2/(2n)!$, for $n = 1, 2, \ldots$, so

$$\begin{aligned} \left|\frac{a_{n+1}}{a_n}\right| &= \frac{((n+1)!)^2}{(2n+2)!} \times \frac{(2n)!}{(n!)^2} \\ &= \frac{(n+1)(n+1)}{(2n+2)(2n+1)}. \end{aligned}$$

Dividing the numerator and denominator by n^2 gives

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{(1+1/n)(1+1/n)}{(2+2/n)(2+1/n)} \to \frac{1}{4} \text{ as } n \to \infty.$$

Hence, by the Ratio Test, the radius of convergence is $R = 4$.

(c) Here $a_n = n + 2^{-n}$, for $n = 0, 1, 2, \ldots$, so

$$\begin{aligned} \left|\frac{a_{n+1}}{a_n}\right| &= \frac{n+1+2^{-n-1}}{n+2^{-n}} \\ &= \frac{1 + 1/n + 1/(n2^{n+1})}{1 + 1/(n2^n)} \to 1 \text{ as } n \to \infty. \end{aligned}$$

Hence, by the Ratio Test, the radius of convergence is $R = 1$.

(d) Here $a_n = n^n$, for $n = 1, 2, \ldots$, so

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{(n+1)^{(n+1)}}{n^n} = (n+1)\left(\frac{n+1}{n}\right)^n \to \infty \text{ as } n \to \infty.$$

Hence, by the Ratio Test, the radius of convergence is $R = 0$; that is, the series converges only for $x = 0$.

Solution to Exercise F64

Applying the Ratio Test with

$$a_n = \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}, \quad \text{for } n \in \mathbb{N},$$

we obtain, for $\alpha \neq 0, 1, 2, \ldots$,

$$\left|\frac{a_{n+1}}{a_n}\right| = \left|\frac{\alpha(\alpha-1)\cdots(\alpha-n)}{\alpha(\alpha-1)\cdots(\alpha-n+1)} \times \frac{n!}{(n+1)!}\right| = \left|\frac{\alpha-n}{n+1}\right| = \left|\frac{(\alpha/n)-1}{1+1/n}\right| \to 1 \text{ as } n \to \infty.$$

Hence the radius of convergence is 1.

Solution to Exercise F65

In each case, we apply Strategy F13.

(a) Here $a_n = n$, for $n = 1, 2, \ldots$.

1. Since

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{n+1}{n} = 1 + 1/n \to 1 \text{ as } n \to \infty,$$

we have $R = 1$, by the Ratio Test. Thus this power series

- converges for $-1 < x < 1$,
- diverges for $x > 1$ and $x < -1$.

2. If $x = 1$, then the power series is

$$\sum_{n=0}^{\infty} n(1)^n = \sum_{n=0}^{\infty} n,$$

which is divergent, by the Non-null Test.

If $x = -1$, then the power series is

$$\sum_{n=0}^{\infty} n(-1)^n = \sum_{n=0}^{\infty} (-1)^n n,$$

which is again divergent, by the Non-null Test.

Hence the interval of convergence is $(-1, 1)$.

(b) Here $a_n = (-1)^n/(n3^n)$, for $n = 1, 2, \ldots$.

1. Since

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{1}{(n+1)3^{n+1}} \times \frac{n3^n}{1} = \frac{1}{(1+1/n)3} \to \frac{1}{3} \text{ as } n \to \infty,$$

we have $R = 3$, by the Ratio Test. Since $a = 5$, this power series

- converges for $2 < x < 8$,
- diverges for $x > 8$ and $x < 2$.

2. If $x = 8$, then the power series is

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n3^n}(8-5)^n = \sum_{n=1}^{\infty} \frac{(-1)^n}{n},$$

which is convergent, by the Alternating Test.

If $x = 2$, then the power series is

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n3^n}(2-5)^n = \sum_{n=1}^{\infty} \frac{1}{n},$$

which is a basic divergent series.

Hence the interval of convergence is $(2, 8]$.

Solution to Exercise F66

(a) We know that, for $x \in \mathbb{R}$,

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + \cdots$$

and

$$e^{-x} = 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \cdots + (-1)^n \frac{x^n}{n!} + \cdots.$$

Hence, by the Sum and Multiple Rules,

$$\sinh x = \tfrac{1}{2}(e^x - e^{-x}) = x + \frac{x^3}{3!} + \cdots + \frac{x^{2n+1}}{(2n+1)!} + \cdots,$$

for $x \in \mathbb{R}$.

(b) We know that, for $|x| < 1$,

$$\log(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \cdots - \frac{x^n}{n} - \cdots$$

and

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots + x^n + \cdots.$$

Hence, by the Sum and Multiple Rules,

$$\begin{aligned}&\log(1-x)+\frac{2}{1-x}\\&=\left(-x-\frac{x^2}{2}-\frac{x^3}{3}-\cdots-\frac{x^n}{n}-\cdots\right)\\&\quad+(2+2x+2x^2+2x^3+\cdots+2x^n+\cdots)\end{aligned}$$

$$=2+x+\frac{3}{2}x^2+\frac{5}{3}x^3+\cdots+\left(2-\frac{1}{n}\right)x^n+\cdots,$$

for $|x|<1$.

Solution to Exercise F67

(a) We know that, for $|x|<1$,

$$\begin{aligned}\log(1+x)&=x-\frac{x^2}{2}+\frac{x^3}{3}-\cdots\\&\quad+(-1)^{n+1}\frac{x^n}{n}+\cdots.\end{aligned}$$

Hence, by the Product Rule,

$$\begin{aligned}&(1+x)\log(1+x)\\&=\left(x-\frac{x^2}{2}+\frac{x^3}{3}-\cdots+(-1)^{n+1}\frac{x^n}{n}+\cdots\right)\\&\quad+\left(x^2-\frac{x^3}{2}+\cdots+(-1)^n\frac{x^n}{n-1}+\cdots\right)\\&=x+\frac{x^2}{2}-\frac{x^3}{6}+\cdots+(-1)^n\frac{x^n}{n(n-1)}+\cdots,\end{aligned}$$

for $|x|<1$.

(b) We know that, for $|x|<1$,

$$\frac{1}{1-x}=1+x+x^2+\cdots+x^n+\cdots.$$

Hence, by the Product Rule,

$$\begin{aligned}\frac{1}{(1-x)^2}&=\left(1+x+x^2+\cdots+x^n+\cdots\right)^2\\&=1+(1+1)x+(1+1+1)x^2+\cdots\\&\quad+(1+1+\cdots+1)x^n+\cdots\\&=1+2x+3x^2+\cdots+(n+1)x^n+\cdots,\end{aligned}$$

for $|x|<1$.

(c) We know that, for $|x|<1$,

$$\frac{1}{1-x}=1+x+x^2+\cdots+x^n+\cdots,$$

and, for $|x|<1$, from Worked Exercise F48, that

$$\frac{1+x}{(1-x)^2}=1+3x+5x^2+\cdots+(2n+1)x^n+\cdots.$$

Hence, by the Product Rule,

$$\begin{aligned}\frac{1+x}{(1-x)^3}&=(1+x+x^2+\cdots+x^n+\cdots)\\&\quad\times(1+3x+5x^2+\cdots+(2n+1)x^n+\cdots)\\&=1+(3+1)x+(5+3+1)x^2+\cdots\\&\quad+((2n+1)+(2n-1)+\cdots+1)x^n+\cdots\\&=1+4x+9x^2+\cdots+(n+1)^2x^n+\cdots,\end{aligned}$$

for $|x|<1$, since $1+3+\cdots+(2n+1)$ is an arithmetic series with sum $(n+1)^2$.

Solution to Exercise F68

(a) We know from Exercise F66(a) and Theorem F65 that, for $x\in\mathbb{R}$,

$$\sinh x=x+\frac{x^3}{3!}+\frac{x^5}{5!}+\cdots+\frac{x^{2n+1}}{(2n+1)!}+\cdots$$

and

$$\sin x=x-\frac{x^3}{3!}+\frac{x^5}{5!}-\cdots+(-1)^n\frac{x^{2n+1}}{(2n+1)!}+\cdots.$$

Hence, by the Sum Rule, for $x\in\mathbb{R}$,

$$\begin{aligned}&\sinh x+\sin x\\&=2x+\frac{2x^5}{5!}+\frac{2x^9}{9!}+\cdots+\frac{2x^{4n+1}}{(4n+1)!}+\cdots.\end{aligned}$$

This Taylor series is valid for all real values of x.

(b) We know that

$$\frac{1}{1+x}=1-x+x^2-\cdots+(-1)^nx^n+\cdots,$$

for $|x|<1$. Replacing x by $2x^2$, we obtain

$$\frac{1}{1+2x^2}=1-2x^2+4x^4-\cdots+(-1)^n2^nx^{2n}+\cdots.$$

This last series converges for $2x^2<1$; that is, for $|x|<1/\sqrt{2}$.

Hence this Taylor series is valid in the interval $(-1/\sqrt{2},1/\sqrt{2})$.

Solution to Exercise F69

We know that

$$e^x=1+x+\frac{x^2}{2!}+\cdots+\frac{x^n}{n!}+\cdots,\quad\text{for } x\in\mathbb{R}.$$

Also, for $|x|<1$, from Exercise F67(b),

$$(1-x)^{-2}=1+2x+3x^2+\cdots+(n+1)x^n+\cdots.$$

Hence, by the Product Rule,

$$\begin{aligned}
&e^x(1-x)^{-2}\\
&= \left(1+x+\tfrac{1}{2}x^2+\cdots\right)(1+2x+3x^2+\cdots)\\
&= 1+(2+1)x+\left(3+2+\tfrac{1}{2}\right)x^2+\cdots\\
&= 1+3x+\tfrac{11}{2}x^2+\cdots, \quad \text{for } |x|<1.
\end{aligned}$$

Solution to Exercise F70

(a) We know from the solution to Exercise F67(b) that, for $|x|<1$,

$$(1-x)^{-2} = 1+2x+3x^2+\cdots+(n+1)x^n+\cdots.$$

Hence, by the Differentiation Rule,

$$2(1-x)^{-3} = 2+6x+\cdots+(n+1)nx^{n-1}+\cdots,$$

for $|x|<1$.

Applying the Multiple Rule, we obtain

$$(1-x)^{-3} = 1+3x+\cdots+\frac{(n+1)n}{2}x^{n-1}+\cdots,$$

for $|x|<1$.

(b) From the table of standard derivatives in the Quick reference section of the module Handbook, we have $f'(x)=\dfrac{1}{1-x^2}$ for $|x|<1$. Also, we know that

$$\frac{1}{1-x} = 1+x+x^2+\cdots+x^n+\cdots, \quad \text{for } |x|<1.$$

Replacing x by x^2, we obtain

$$f'(x) = 1+x^2+x^4+\cdots+x^{2n}+\cdots, \quad \text{for } |x|<1.$$

Thus, by the Integration Rule, the Taylor series at 0 for f is

$$f(x) = c+x+\frac{x^3}{3}+\frac{x^5}{5}+\cdots+\frac{x^{2n+1}}{2n+1}+\cdots,$$

for $|x|<1$. Since $f(0)=0$, it follows that $c=0$.

Hence

$$\tanh^{-1}x = x+\frac{x^3}{3}+\frac{x^5}{5}+\cdots+\frac{x^{2n+1}}{2n+1}+\cdots,$$

for $|x|<1$.

Solution to Exercise F71

Since

$$e^x = 1+x+\frac{x^2}{2!}+\cdots+\frac{x^n}{n!}+\cdots, \quad \text{for } x\in\mathbb{R},$$

we have

$$e^{-x^2} = 1-x^2+\frac{x^4}{2!}-\cdots+(-1)^n\frac{x^{2n}}{n!}+\cdots,$$

for $x\in\mathbb{R}$. It follows from the Integration Rule that

$$\begin{aligned}
&\int_0^1 e^{-x^2}\,dx\\
&= \left[x-\frac{x^3}{3}+\frac{x^5}{5\times 2!}-\cdots+(-1)^n\frac{x^{2n+1}}{(2n+1)n!}+\cdots\right]_0^1\\
&= 1-\frac{1}{3}+\frac{1}{10}-\cdots+\frac{(-1)^n}{(2n+1)n!}+\cdots.
\end{aligned}$$

Solution to Exercise F72

We know that

$$\frac{d}{dx}\log(1+x) = \frac{1}{1+x},$$

and that, for $|x|<1$,

$$\frac{1}{1+x} = 1-x+x^2-\cdots+(-1)^nx^n+\cdots.$$

Hence, by the Integration Rule,

$$\log(1+x) = c+x-\frac{x^2}{2}+\cdots+(-1)^n\frac{x^{n+1}}{n+1}+\cdots,$$

for $|x|<1$, where c is a constant.

On substituting $x=0$, we find that $c=0$. Hence

$$\log(1+x) = x-\frac{x^2}{2}+\cdots+(-1)^{n+1}\frac{x^n}{n}+\cdots,$$

for $|x|<1$. (Note that $(-1)^{n+1}=(-1)^{n-1}$.)

Solution to Exercise F73

By the General Binomial Theorem,

$$(1+4x)^{-1/3} = \sum_{n=0}^{\infty}\binom{-\frac{1}{3}}{n}(4x)^n, \quad \text{for } |4x|<1,$$

where $\dbinom{-\frac{1}{3}}{0}=1$ and

$$\binom{-\frac{1}{3}}{n} = \frac{\left(-\frac{1}{3}\right)\left(-\frac{4}{3}\right)\left(-\frac{7}{3}\right)\cdots\left(-\frac{1}{3}-n+1\right)}{n!}, \quad \text{for } n\in\mathbb{N}.$$

Hence

$$\begin{aligned}
(1+4x)^{-1/3} &= 1+\frac{\left(-\frac{1}{3}\right)}{1}4x+\frac{\left(-\frac{1}{3}\right)\left(-\frac{4}{3}\right)}{2!}(4x)^2+\cdots\\
&= 1-\tfrac{4}{3}x+\tfrac{32}{9}x^2-\cdots, \quad \text{for } |x|<\tfrac{1}{4}.
\end{aligned}$$

Acknowledgements

Grateful acknowledgement is made to the following sources.

Cover image: © Mark Owen

Unit F1

Teiji Takagi (Subsection 3.2): Shigeru Tamura / Wikipedia

Karl Weierstrass (Subsection 3.3): Bygone Collection / Alamy Stock Photo

Unit F2

Charles Hermite (Subsection 3.2): By kind permission of the London Mathematical Society

The Marquis de l'Hôpital (Subsection 5.2): Wellcome Collection. This file is licensed under the Creative Commons Attribution Licence http://creativecommons.org/licenses/by/3.0/

Unit F3

Georg Friedrich Bernhard Riemann (Subsection 1.2): Familienarchiv Thomas Schilling, Wikipedia

Henri Léon Lebesgue (Subsection 1.3): http://www.math.ru/history/people/Lebesgue, Wikipedia

John Wallis (Subsection 3.2): Source Wikipedia. This file is licensed under the Creative Commons Attribution-Share Alike Licence http://creativecommons.org/licenses/by-sa/3.0/

Colin Maclaurin (Subsection 3.3): NYPL / Universal Images Group. Britannica Image Quest

Unit F4

Brook Taylor (Subsection 2.1): National Portrait Gallery / Universal Images Group. Britannica Image Quest

Every effort has been made to contact copyright holders. If any have been inadvertently overlooked the publishers will be pleased to make the necessary arrangements at the first opportunity.

Index